

SEARCH REQUEST FORM  
Scientific and Technical Information Center

Access DB#

57687

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

\*\*\*\*\*  
STAFF USE ONLY

Searcher: Sheppard

Searcher Phone #: 308-4499

Searcher Location: \_\_\_\_\_

Date Searcher Picked Up: \_\_\_\_\_

Date Completed: \_\_\_\_\_

Searcher Prep & Review Time: \_\_\_\_\_

Clencal Prep Time: \_\_\_\_\_

Online Time: \_\_\_\_\_

Type of Search

NA Sequence (#) \_\_\_\_\_

AA Sequence (#) \_\_\_\_\_

Structure (#) \_\_\_\_\_

Bibliographic \_\_\_\_\_

Litigation \_\_\_\_\_

Fulltext \_\_\_\_\_

Patent Family \_\_\_\_\_

Other \_\_\_\_\_

Vendors and cost where applicable

STN \_\_\_\_\_

Dialog \_\_\_\_\_

Questel/Orbit \_\_\_\_\_

Dr. Link \_\_\_\_\_

Lexis/Nexis \_\_\_\_\_

Sequence Systems \_\_\_\_\_

WWW/Internet \_\_\_\_\_

Other (specify) \_\_\_\_\_



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 06:50:11 : Search time 4812.74 Seconds

(without alignments)  
13528.937 Million cell updates/sec

Title: US-09-502-426-1

Perfect score: 6888

Sequence: 1 atgttggtattatattgttg.....cctagtctagccaagcct 6888

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 0%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: gb\_est41:\*  
42: gb\_est42:\*  
43: gb\_est43:\*  
44: gb\_est44:\*  
45: gb\_est45:\*  
46: gb\_est46:\*  
47: gb\_est47:\*  
48: gb\_est48:\*  
49: gb\_est49:\*  
50: gb\_est50:\*  
51: gb\_est51:\*  
52: gb\_est52:\*  
53: gb\_est53:\*  
54: gb\_est54:\*  
55: gb\_est55:\*  
56: gb\_est56:\*  
57: gb\_est57:\*  
58: gb\_est58:\*  
59: gb\_est59:\*  
60: gb\_est60:\*  
61: gb\_est61:\*  
62: gb\_est62:\*  
63: gb\_est63:\*  
64: gb\_est64:\*  
65: gb\_est65:\*  
66: gb\_est66:\*  
67: gb\_est67:\*  
68: gb\_est68:\*  
69: gb\_est69:\*  
70: gb\_est70:\*  
71: gb\_est71:\*  
72: gb\_est72:\*  
73: gb\_est73:\*  
74: gb\_est74:\*  
75: gb\_est75:\*  
76: gb\_est76:\*  
77: gb\_est77:\*  
78: gb\_est78:\*  
79: gb\_est79:\*  
80: gb\_est80:\*  
81: gb\_est81:\*  
82: gb\_est82:\*  
83: gb\_est83:\*  
84: gb\_est84:\*  
85: gb\_est85:\*  
86: gb\_est86:\*  
87: gb\_est87:\*  
88: gb\_est88:\*  
89: gb\_est89:\*  
90: gb\_est90:\*  
91: gb\_est91:\*  
92: gb\_est92:\*  
93: gb\_est93:\*  
94: gb\_est94:\*  
95: gb\_est95:\*  
96: gb\_est96:\*  
97: gb\_est97:\*  
98: gb\_est98:\*  
99: gb\_est99:\*  
100: gb\_est100:\*  
101: gb\_est101:\*  
102: gb\_est102:\*  
103: gb\_est103:\*  
104: gb\_est104:\*  
105: gb\_est105:\*  
106: gb\_est106:\*  
107: gb\_est107:\*  
108: gb\_est108:\*  
109: gb\_est109:\*  
110: gb\_est110:\*  
111: gb\_est111:\*  
112: gb\_est112:\*  
113: gb\_est113:\*  
114: gb\_est114:\*  
115: gb\_est115:\*  
116: gb\_est116:\*

44: em\_esthum10:\*  
45: em\_esthum11:\*  
46: em\_esthum12:\*  
47: em\_esthum13:\*  
48: em\_esthum14:\*  
49: em\_esthum15:\*  
50: em\_esthum16:\*  
51: em\_esthum17:\*  
52: em\_esthum18:\*  
53: em\_esthum19:\*  
54: em\_esthum20:\*  
55: em\_esthum21:\*  
56: em\_esthum22:\*  
57: em\_esthum23:\*  
58: em\_esthum24:\*  
59: em\_esthum25:\*  
60: em\_esthum26:\*  
61: em\_esthum27:\*  
62: em\_esthum28:\*  
63: em\_esthum29:\*  
64: em\_esthum30:\*  
65: em\_esthum31:\*  
66: em\_esthum32:\*  
67: em\_esthum33:\*  
68: em\_esthum34:\*  
69: em\_esthum35:\*  
70: em\_esthum36:\*  
71: em\_esthum37:\*  
72: em\_esthum38:\*  
73: em\_esthum39:\*  
74: em\_esthum40:\*  
75: em\_esthum41:\*  
76: em\_esthum42:\*  
77: em\_esthum43:\*  
78: em\_esthum44:\*  
79: em\_esthum45:\*  
80: em\_esthum46:\*  
81: em\_esthum47:\*  
82: em\_esthum48:\*  
83: em\_esthum49:\*  
84: em\_esthum50:\*  
85: em\_esthum51:\*  
86: em\_esthum52:\*  
87: em\_esthum53:\*  
88: em\_esthum54:\*  
89: em\_esthum55:\*  
90: em\_esthum56:\*  
91: em\_esthum57:\*  
92: em\_esthum58:\*  
93: em\_esthum59:\*  
94: em\_esthum60:\*  
95: em\_esthum61:\*  
96: em\_esthum62:\*  
97: em\_esthum63:\*  
98: em\_esthum64:\*  
99: em\_esthum65:\*  
100: em\_esthum66:\*  
101: em\_esthum67:\*  
102: em\_esthum68:\*  
103: em\_esthum69:\*  
104: em\_esthum70:\*  
105: em\_esthum71:\*  
106: em\_esthum72:\*  
107: em\_esthum73:\*  
108: em\_esthum74:\*  
109: em\_esthum75:\*  
110: em\_esthum76:\*  
111: em\_esthum77:\*  
112: em\_esthum78:\*  
113: em\_esthum79:\*  
114: em\_esthum80:\*  
115: em\_esthum81:\*  
116: em\_esthum82:\*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*

```

```

190: gb_est110:*
191: gb_est111:*
192: gb_est112:*
193: gb_est113:*
194: gb_est114:*
195: gb_est115:*
196: gb_est116:*
197: gb_est117:*
198: gb_est118:*
199: gb_est119:*
200: gb_est120:*
201: gb_est121:*
202: gb_est122:*
203: gb_est123:*
204: gb_est124:*
205: gb_est125:*
206: gb_est126:*
207: gb_est127:*
208: gb_est128:*
209: gb_est129:*
210: gb_est130:*
211: gb_est131:*
212: gb_est132:*
213: gb_est133:*
214: gb_est134:*
215: gb_est135:*
216: gb_est136:*
217: gb_est137:*
218: gb_est138:*
219: gb_est139:*
220: gb_est140:*
221: gb_est141:*
222: gb_est142:*
223: gb_est143:*
224: gb_est144:*
225: gb_est145:*
226: gb_est146:*
227: gb_est147:*
228: gb_est148:*
229: gb_est149:*
230: gb_est150:*
231: gb_est151:*
232: gb_est152:*
233: gb_est153:*
234: gb_est154:*
235: gb_est155:*
236: gb_est156:*
237: gb_est157:*
238: gb_est158:*
239: gb_est159:*
240: gb_est160:*
241: gb_est161:*
242: gb_est162:*
243: gb_est163:*
244: gb_est164:*
245: gb_est165:*
246: gb_est166:*
247: gb_est167:*
248: gb_est168:*
249: gb_est169:*
250: gb_est170:*
251: gb_est171:*
252: gb_est172:*
253: gb_est173:*
254: gb_est174:*
255: gb_est175:*
256: gb_est176:*
257: gb_est177:*
258: gb_est178:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.





BE315722  
LOCUS BE315722 642 bp mRNA EST 14-JUL-2000  
DEFINITION NF025D1LF1090 Developing leaf Medicago truncatula cDNA clone  
NE025D1LF 5', mRNA sequence.  
ACCESSION BE315722  
VERSION BE315722.1 GI:9189499  
KEYWORDS EST.  
SOURCE barrel medic.  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.  
1 (bases 1 to 642)  
Flóres-H.R., Inman,J.T., Weller,J.W. and May,G.D.  
AUTHORS  
Expressed Sequence Tags from the Samuel Roberts Noble Foundation  
Medicago truncatula leaf library  
JOURNAL Unpublished (2000)  
COMMENT Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Medicago Genome Initiative accession: MGI:S:16068  
Insert Length: 642 Std Error: 0.00  
Plate: 025 row: D column: 11  
Seq primer: TCACACGAGAAACAGCTATGAC.  
FEATURES  
source  
1..642  
/organism="Medicago truncatula"  
/db\_xref="taxon:3880"  
/clone="NF025D1LF"  
/clone\_lib="Developing leaf"  
/tissue\_type="leaf"  
/dev\_stage="Pooled developmental"  
/note="Vector: Lambda Zap; Contains a mixture of very  
young, developing, mature and senescing leaves."  
BASE COUNT 210 a 110 c 109 g 213 t  
ORIGIN  
Query Match 2.7%; Score 186.8; DB 166; Length 642;  
Best Local Similarity 76.2%; Pred. No. 9.5e-22;  
Matches 230; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
OY 3502 aggtatgtaagatatagatgaactgtttggagaacacgacgtatcagctgat 3561  
|||||  
DB 322 AGGTATGCTAAATTTTCAAAATCAAACTTTTGGTGAGACCAACATAGTTTCAAGCAT 361  
OY 3562 gctggacttaatatcatatatacaaaacgaagagctcttgaatgtatctc 3621  
|||||  
DB 382 GCAGATGTAATAGCTCATATATCAAAATGAAGGAAATATTATTGATGTCAGTTATCCT 441  
OY 3622 agaagataagtggtgactcttggaagaaatggtcgtatgcttcttggtagcagatc 3681  
|||||  
DB 442 AGTAACTTTGTTGGAATATCTGGAAATGCTCATGTTGTTGGTGTGACATGCAT 501  
OY 3682 agagatatggaagtactcttaacttcttaagtcagcagctcttaactattc 3741  
|||||  
DB 502 AGAGACTGCTGTAATATATACGTAATTTTGTGTCATCTAGGCTTAGACACATCTA 561  
OY 3742 cttaaagatctgagaaacactcttcttcttgaatcttgacacaaactatc 3801  
|||||  
DB 562 TTGAAAGAGTTGGAACATACCCGCTTGTTTAAGTCTTGGAAAGAAACCATCAT 621  
OY 3802 tct 3803  
||  
DB 622 TT 623  
RESULT 3

BF324723  
LOCUS BF324723 667 bp mRNA EST 21-NOV-2000  
DEFINITION su14f03.y1 Gm-cl066 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl066-1781 5', similar to TR:064989 064989 STEROID  
22-ALPHA-HYDROXYLASE. [1] ;, mRNA sequence.  
ACCESSION BF324723  
VERSION BF324723.1 GI:11274370  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.  
1 (bases 1 to 667)  
Shoemaker,R., Kelm,P., Vodkin,L., Erepding,J., Corvelli,V., Rhanna  
A., Bolla,B., Merrin,M., Hillier,L., Kueba,T., Martin,J., Beck,C.,  
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers  
Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk  
R., Rilter,E., Kohn,S., Shin,T., Jackson,T., Cardenas,M., McCann  
R., Waterston,R. and Willson,R.  
AUTHORS  
Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Genome Systems, Inc. 4633 World  
Parkway Circle St. Louis, Missouri 63134 For further information  
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
427-3324 or contact: clones@genomesystems.com or  
info@genomesystems.com web site: www.genomesystems.com  
High quality sequence stop: 420.  
FEATURES  
source  
1..667  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl066-1781"  
/clone\_lib="Gm-cl066"  
/tissue\_type="leaf and shoot tip, salt stressed, 2 week  
old seedling"  
/lab\_host="DH10B"  
/note="Vector: pBluescript II SK+; Site\_1: EcoRI; Site\_2:  
XhoI; The cDNA library was constructed from mRNA isolated  
from unexpanded leaves and the shoot tips of 2 week old  
seedling from the cultivar Williams. The 2 week old  
seedlings were salt stressed in a solution of 500mM NaCl  
for 3 days prior to harvesting. Complementary DNA was  
synthesized from mRNA using a primer consisting of a  
poly(dT) sequence with a XhoI restriction site. EcoRI  
adapters were ligated to the blunt-ended cDNA fragments  
followed by XhoI digestion. The cDNA fragments were  
directionally cloned into the EcoRI-XhoI restriction site  
of the pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (GibcoBRL). This library  
was constructed in the laboratory of Dr. Randy  
Shoemaker."  
BASE COUNT 194 a 155 c 141 g 177 t  
ORIGIN  
Query Match 2.6%; Score 178.6; DB 147; Length 667;  
Best Local Similarity 71.4%; Pred. No. 2.3e-20;  
Matches 235; Conservative 0; Mismatches 94; Indels 0; Gaps 0;  
OY 3502 aggtatgtaagatatagatgaactgtttggagaacacgacgtatcagctgat 3561  
|||||  
DB 218 AGGTATGCTAAATTTTCAAAATCAAACTTTTGGTGAGACCAACATAGTTTCAAGCAT 277  
OY 3562 gctggacttaatatcatatatacaaaacgaagagctcttgaatgtatctc 3621  
|||||

Db 278 GCAGACTCAACAGTTCATTCTACAAAACGAGGAATTTGTCAGTGCAGCTATCCT 337  
 OY 3622 agaagtagtgagtgatctctgggaatgfcagatgcttctgttggtgacatgat 3681  
 Db 338 AGAAGCATCGGTGGAATCTAGGAATATGGTCATGTTGGCTTAGTGGACATGCAT 397  
 OY 3682 agagatagagaatctatctgcttaactcttaagtcacagcagctcttagaactatctca 3741  
 Db 398 AGAAGCATCGGTGGAATCTAGGAATATGGTCATGTTGGCTTAGTGGACATGCAT 457  
 OY 3742 cttaagatgttgtagagacatacttggcttcttgatcttgagcaacaactatct 3801  
 Db 458 TTGAAGAGGCGGAGGAGGCTATCCCTCTTGCTGAACCTCTTGAGCCAGATTCATA 517  
 OY 3802 ttctcgtctgaagcagcgccaaagct 3830  
 Db 518 TTCTAGCCCAAGATGAGCTAAGAGATT 546

RESULT 4  
 LOCUS BF050501 533 bp mRNA EST 16-OCT-2000  
 DEFINITION EST435639 tomato developing/immature green fruit Lycopersicon  
 esculentum cDNA clone cLEM18622 5' sequence, mRNA sequence.  
 ACCESSION BF050501  
 VERSION BF050501.1 GI:10804397  
 KEYWORDS EST.  
 SOURCE tomato.  
 ORGANISM Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;  
 Lycopersicon.  
 REFERENCE 1 (bases 1 to 533)  
 AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
 Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.D., Romling,C.M.,  
 Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley  
 S.D.  
 TITLE Generation of ESTs from tomato fruit tissue, immature green  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: David Frisch  
 Clemson University Genomics Institute  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 4366  
 Fax: 864 656 4293  
 Email: dfrisch@CLEMSON.EDU.

## FEATURES

SOURCE  
 Location/Qualifiers  
 1..533  
 /organism="Lycopersicon esculentum"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="cLEM18622"  
 /clone\_lib="tomato developing/immature green fruit"  
 /tissue\_type="fruit"  
 /dev\_stage="immature green (5-35 days post-anthesis)"  
 /lab\_host="SOLR"  
 /note="Vector: pBluescriptSKMunadpt. Site\_1: EcoRI;  
 Site\_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and  
 harvested at 7 day intervals through 35 dpa. Equal masses  
 of tissue from each stage were combined (including seeds  
 and locules) prior to mRNA isolation."

BASE COUNT 172 a 75 c 111 g 175 t  
 ORIGIN

Query Match 2.5%; Score 171.4; DB 143; Length 533;  
 Best Local Similarity 73.1%; Pred. No. 3.7e-19;  
 Matches 220; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

OY 3497 ttacagatagtgatagatatagatgaactgttggagaaacagatcgtatcag 3556  
 Db 233 TTCTAGGTATGGAAATTTTCAAGTCAAAATTTGTTGGAGAGCCACAAATAGTTTCAG 292

OY 3557 ctgatgcggaacttaaatagatctcatatcaaaaggaaggctcttgaatgatt 3616  
 Db 293 CAGATGCAGGCGCTTAACAGATATCATTTGCGAAGAAAGAGGAGATTATTGAGTGAATT 352  
 OY 3617 atccctaagaatagtgatgcttctgggaatgctgcatgcttcttctgttgaca 3676  
 Db 353 ATCCAGAAAGATATAGTGTGATGATCTGTAATAGTGTATGTTAGTTGACACAAA 412  
 OY 3677 tgcataagatagagaatctctgcttaactcttaagtcacagcagctcttagaacta 3736  
 Db 413 TGCAATAGAGATATGAGATGATCTCTGAAATTTTTCAGCAATGCTAGCTAAGATATC 472  
 OY 3737 ttctactaaagatgttagagacatacttggcttcttgatcttcttgagcaacaact 3796  
 Db 473 AACTTTAAGTGAAGTGAAGATGATGATCTGCTGTTGCTTGCTTGGAACAGGATT 532  
 OY 3797 c 3797  
 Db 533 c 533

RESULT 5  
 LOCUS BG316131 522 bp mRNA EST 26-FEB-2001  
 DEFINITION BG316131  
 ID: Gm-cl032-2967 5' similar to TR:064989 064989 STEROID  
 22-ALPHA-HYDROXYLASE. [I] ; mRNA sequence.  
 ACCESSION BG316131  
 VERSION BG316131.1 GI:13125561  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.  
 REFERENCE 1 (bases 1 to 522)  
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Expellding,J., Corvelli,V., Khanna  
 A., Bolla,B., Marra,M., Hillier,L., Kucab,J., Martin,J., Beck,C.,  
 Wylie,T., Underwood,K., Scepcoe,M., Theising,B., Allen,M., Bowers  
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk  
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann  
 R., Waterston,R. and Wilson,R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu

FEATURES  
 SOURCE  
 Location/Qualifiers  
 1..522  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl032-2967"  
 /clone\_lib="Gm-cl032"  
 /tissue\_type="Cotyledons of 8-day-old 'Williams'  
 seedlings"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from cotyledons of 8-day-old 'Williams' seedlings which  
 were propagated on paper towels with distilled water for  
 3 days (etiolated), then greenhouse grown for 5 days in  
 potting soil. The cotyledons were flash-frozen in liquid



Fax: 864 656 4293  
 Email: rwing@clemons.edu  
 Seq primer: GGAAACAGCTATGACCATG  
 Class: BAC ends  
 High quality sequence stop: 191.  
 Location/Qualifiers  
 1. 589

FEATURES  
 source  
 /organism="Oryza sativa"  
 /strain="Japonica"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:4530"  
 /clone="dbx0078N12r"  
 /clone\_lib="CUGI Rice BAC Library"  
 /issue\_type="leaf"  
 /lab\_host="E. coli DH10B"  
 /note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 Kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."  
 BASE COUNT 112 a 181 c 181 g 114 t 1 others  
 ORIGIN

Query Match 2.1% Score 146; DB 232; Length 589;  
 Best Local Similarity 57.9%; Pred. NO. 6.7e-15;  
 Matches 288; Conservative 0; Mismatches 191; Indels 18; Gaps 1;  
 Oy 3499 tacaggtatgttaagatatagatcgaactgttttgagagacaacagtcgtacagt 3558  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 85 TGCAAGTACGGGGAAGATATACCGGTCGAGCCGTTCCGGGAGCGAGCGGTGTCGCG 144  
 Oy 3559 gatctgacttaataagatcatattacaagaagaagcctcttgatgtatgtat 3618  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 145 GACGCGGGCTCAACCGGTACATCCTGCAGACGAGGAGCGTGTGAGTGCAGCTAC 204  
 Oy 3619 cctagaagataggttggaattcttggaagaatgcatgctgtgtctgttgatgacatg 3678  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 205 CCGGCGACATCGCGGCGATCTGGGCAAGTGTCATGCTGTCCTGTCGGGAGCCG 264  
 Oy 3679 catagagatagagaagatctcgcttaactcttaagtlcagcagcgtcttagaactat 3738  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 265 CACGCGGAGATGCGGCGCATCTCCCTCAACTTCTCCTCCGTCGCGCGCCGCTC 324  
 Oy 3739 ctaactaaagatgttgagacataacttgttcttgattcttgcaacaacaactct 3798  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 325 CTCTCCCGCGGAGTGCAGCGCACCCCTCCTCGTCTCCGCGCTGCCCTTCCTCC 384  
 Oy 3799 atttctctgctcaagaagagcaaaaagt-----tttattctt 3840  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 385 ACCTTCTTCGCTCAACACCAAGCCAGAGGATACACACATCGTCTCCATTAATGATG 444  
 Oy 3841 acccttatttctgctaaatttttctttagaactcttagagtttcaactttttt 3900  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 445 AGCATGCAATATATCAATAGCTCGCTAGCGATCAATATCAATTAAGCGCGTTTCG 504  
 Oy 3901 ttttaattgaacagtttaacttcaatcgaagcagataatagatgatgactcctg 3960  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 505 TGCGTGTCTGAGTTCACTGATGAGCGAAGAAATATGACCATGAGACCCCG 564

Oy 3961 agaagaagaacagagc 3977  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 565 GGGAGAGAGAGAGAGAGC 581

RESULT 8  
 CENS00EVL/c 1101 bp DNA GSS 04-JUN-1999  
 LOCUS  
 DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC: BACR29B23 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
 ACCESSION AL069706  
 VERSION AL069706.1 GI:4949849  
 KEYWORDS GSS.  
 SOURCE  
 ORGANISM  
 Drosophila melanogaster  
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 1 (bases 1 to 1101)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)  
 Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoko Osoegawa and Aaron Mammoser at Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACRC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).  
 Location/Qualifiers

FEATURES  
 source  
 /organism="Drosophila melanogaster"  
 /db\_xref="taxon:7227"  
 /clone\_lib="RPc1-98"  
 /clone="BACR29B23"  
 /note="end : 57"  
 BASE COUNT 419 a 91 c 60 g 299 t 232 others  
 ORIGIN  
 Query Match 1.6% Score 113.2; DB 219; Length 1101;  
 Best Local Similarity 35.8%; Pred. NO. 2.1e-09;  
 Matches 232; Conservative 118; Mismatches 297; Indels 1; Gaps 1;  
 Oy 1625 tagatgagatattccaatcaacttggatgattaaacttaatttcttct 1684  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1100 TWTATWTTWTTWTTWTTWTTATATATATATATATATATATATATATATATATAT 1041  
 Oy 1685 acggtataataatcaatcagagtaaaaaagttgtcttatttcgcgatgcatga 1744  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 1040 WATATATATWTTWTTAAATATATATATATATATATATATATATATATATATATAA 981  
 Oy 1745 aggaataacctaagacttaatttttgaagaatgaacccttatactacagatgaact 1804  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 980 WTATWATWTTAAATWATWATWATWATWATWATWATWATWATWATWATWATWAT 921  
 Oy 1805 accgaatgttttctgccaataagcagcctctcaactgtgataagcaatltttctg 1864  
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 Db 920 TTTWTTTATWTTTATWTTTATWTTWMAAAMWMAAMWMAAMWMAAMWMAAMW 861  
 Oy 1865 caaatataattgaatcaatgctactatcaatagaagaacagctgattatcat 1924

[illegible]

Query Match Similarity 1.6%; Score 110.8; DB 219; Length 1101;  
Best Local Similarity 34.5%; Pred. No. 5.3e-09;  
Matches 223; Conservative 128; Mismatches 290; Indels 5; Gaps 1;

OY 1513 taaccatcgcgacattcaaatctgtcttcctaatacgcatccctaacaatatattatgttag 1572  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 461 TTTTMMNNMNMNNAATTTWNAANAAVMTTAATMAATAAAAAAAMMMMMATTTTTTMMMNNTWA 520  
OY 1573 ttggaattcaagaagcgaactcgttaacatacaatatattatagactagtatcgt 1632  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 521 TTWTWTMMNMTTATTAATAAAAAMAAATAATTTAAAMMAATTAATTAAMAATTTTMAAMW 580  
OY 1633 gattaaccaaataactcgttggttggtaacctaaatctgtttctccacggtata 1692  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 581 TATATTAATTTATTAATATMTATTAATATTAATAAAATATTTTWTAAAAATTTTAAATPA 640  
OY 1693 aaataaacatcgcagggtaaaanaagtttgtctattcttcgcgacgcgaaggataa 1752  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 641 TTTTATTTAMTTATTAATATPAAWMTAAWTMMWTATTAATTAATTTTMAAATPMAAAAAAAA 700  
OY 1753 cctaatagccttaatttttggaaatgtacccctttactccatagataatcacgtatg 1812  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 701 AAAAAAMATPAAANAATTAATMAATPAAATTTAAAAAATAAAMAAATPMAATPAAATPMA 760  
OY 1813 ttttgttgccttacagcagcctcacacacgtgatagtaacttttcgcgaatat 1872  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 761 TATATATTTTATMMNMTMMNMTTANA-----TANATTAAMAMAMATTAATATMA 815  
OY 1873 aaattagggaatccaatgctactcataatagaagaacacgcgtagtatcatcttaatt 1932  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 816 WATMAAANAAMATATMMWTATATAMNATAMNATAAAMAAAAATTAATATPAAATMAA 875  
OY 1933 aaagcacaatttttggaaaaatgctataattctcaacaatatataaataatgtgct 1992  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 876 AAAATAMAMTWTMTTTTMMAMMATPATPAAMAMTATPAAAAAAMAAAAAATTAAMA 935  
OY 1993 atcaatgttcccatcgtctcctaataatttttcttattatgtactataacattat 2052  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 936 WMTATATTTTWTATTAAMTWATMTATMMATTTMMATMTTATTTATTTATATATAT 995  
OY 2053 gaaccaataatgctgtgtaaltcaaatcatctcaatatatttttgaatctacaat 2112  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 996 WTAMTAMATATTTATTTAAAMWTATATTTTAAAMWTAAATATATATATATMTAMTAT 1055  
OY 2113 tatataatttagtcaataacatgcatagaagttccaaaaaaa 2158  
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : :  
DB 1056 AAAMAAATTAATTAATATTAAMATMAAAMAAAAAMAAAMATPMAA 1101

RESULT 10  
CN50021J/C  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

CNS0021J 1101 bp DNA GSS 03-JUN-1999  
Drosophila melanogaster genome survey sequence TEM3 end of BAC #  
BACR03N11 of RPCI-98 library from Drosophila melanogaster (fruit  
fly), genomic survey sequence.  
AL061936  
AL061936.1 GI:4940214  
GSS.  
fruit fly.  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
Muscomorpha; Ephyroidae; Drosophilidae; Drosophila.  
1 (bases 1 to 1101)  
Genoscope.  
Submitted Submission  
Direct Submission  
BP 191 J1006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a  
collaboration with the Berkeley Drosophila Genome Project (BDGP).  
The BDGP is constructing a physical map of the Drosophila





AM398669	LOCUS	AM398669	615 bp	mRNA	EST	07-FEB-2000
DEFINITION	ESR309169 L. pennellii trichome, Cornell University Lycopersicon pennellii cDNA clone cLPT4K22 5', mRNA sequence.					
ACCESSION	AM398669					
VERSION	AM398669.1					
KEYWORDS	EST.					
SOURCE	Lycopersicon pennellii.					
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;					
REFERENCE	1 (bases 1 to 615)					
AUTHORS	Alcala,J., Vrebalov,J., White,R., Materer,A.L., Lakey,M., Holt,I.E., Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Niernman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni,J.					
TITLE	Generation of ESTs from wild tomato (Lycopersicon pennellii) trichomes					
JOURNAL	unpublished (1999)					
COMMENT	Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfirsch@CLEMSON.EDU					
FEATURES	5 prime sequence.					
source	Location/Qualifiers					
	1..615					
	/organism="Lycopersicon pennellii"					
	/db_xref="taxon:28526"					
	/clone="CLPT4K22"					
	/clone_11b="L. pennellii trichome, Cornell University"					
	/tissue="trichome"					
	/dev_stage="mixed stages"					
	/lab_host="SOLR"					
	/note="vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells highly enriched for trichomes, with minor contamination by other types of leaf cells."					
BASE COUNT	180 a 111 c 100 g 224 t					
ORIGIN						
Query Match	1.5%; Score 105; DB 115; Length 615;					
Best Local Similarity	57.0%; Pred. No. 5.1e-08;					
Matches 192; Conservative	0; Mismatches 145; Indels 0; Gaps 0;					
QY	3496 ttcttaagatagcgaagatatatagatgcagcactggtttggaacacagatcgtatca	3555				
DB	265 tctttccttaggtatgggaagactgttcaagtcacattttttttccccaacagtggtca	324				
QY	3556 gctgtagctgcgactaatagatcatatcacaaacgaagaagcgtcttgaagctagt	3615				
DB	325 ttgtgaccaaagaccttaattacttcattatattacaaatgaagataagttatttcagtgagt	384				
QY	3616 tatctagaagctatagctgggaattcttgggaatgctgcagatgcttctgttggtagc	3675				
DB	385 tatccaaagcccaattcattcagcttactgttgccaaagtttcatttccttgcttgctgtgac	444				
QY	3676 atgcagtagagatagagaagatgctgcgtcttaactcttaagtcagaacgctctagaact	3735				
DB	445 acacataaaaggcttagcgaatttttcattattatcattatcagcacacatttaagcttaacct	504				
QY	3736 attactactaaagatgcttgagagacatactgttcttctgctgattcttggagaacaaac	3795				
DB	505 gagtttatattaatgattgtgaacatttagacacttgagattcttcattcatatgaaaagatpaa	564				
QY	3796 tctattctctgcctcaagaagagccaaaagctt 3832					

DB	565	CATCAAGTCACGACTGGGAAGAGCCAAAGAAGTTT	601
RESULT	13		
LOCUS	BF052051	191 bp	mRNA EST 16-OCT-2000
DEFINITION	EST437298 tomato developing/immature green fruit Lycopersicon esculentum cDNA clone cLEM25M4 5' sequence, mRNA sequence.		
ACCESSION	BF052051		
VERSION	BF052051.1	GI:10805947	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		
Eukaryota; Viridiplantae;	Streptophyta; Embryophyta; Tracheophyta;		
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;			
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;			
Lycopersicon.			
REFERENCE	1 (bases 1 to 191)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Roming,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tankeley S.D.		
TITLE	Generation of ESTs from tomato fruit tissue, immature green unpublished (2000)		
JOURNAL	Contact: David Frisch		
COMMENT	Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU. Location/Dualifiers		
FEATURES	source		
	1..191		
	/organism="Lycopersicon esculentum"		
	/cultivar="TA496"		
	/db_xref="taxon:4081"		
	/clone="cLEM25M4"		
	/clone_1lb="tomato developing/immature green fruit"		
	/tissue_type="fruit"		
	/dev_stage="Immature green (5-35 days post-anthesis)"		
	/lab_host="SOLR"		
	/note="Vector: pBluescriptSKmCuadapt; Site:1: EcoRI; Site:2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and harvested at 7 day intervals through 35 dpa. Equal masses of tissue from each stage were combined (including seeds and locules) prior to mRNA isolation."		
BASE COUNT	65 a 19 c 46 g 61 t		
ORIGIN			
Query Match	1.5%:	Score 103.6:	DB 143: Length 191:
Best Local Similarity	71.6%:	Pred. No. 8.9e-08:	
Matches 136: Conservative	0: Mismatches 54: Indels	0: Gaps	0:
OY	3574	agatcatataaacaaaagaaggactttgaagttagtacctagaagataaggt	3633
Db	2	AGATCATATTGCACAAGAACGAGGAGATTATTGTAGCTAATTATCCAAGAAAGTATAGCT	61
OY	3634	gggattcttggaagaatggtcgatgcgtcttcttgttgtagacagcataagatataga	3693
Db	62	GCGATACCTTGGTAAATGGCTATGTTAAGTCAAGTTGGACAAATGCCATAAGATATGAGG	121
OY	3694	agatcatcgcttaaccttctaagtcacgacgctcttagaactatctactaaagatgtt	3753
Db	122	ATGATTTTCTCGAATTTTGTAGCAATGCTATGCTAAGGATCAACATTTTAAAGTCAAGTT	181
OY	3754	gagagacata 3763	
Db	182	GAAAGCATAT 191	
RESULT	14		
	FW615836		



LOCUS	AM615836	574 bp	mRNA	EST	24-MAR-2000
DEFINITION	E93325334 tomato flower buds cDNA clone CTOAL7E20 5', mRNA sequence.				
ACCESSION	E93325334				
VERSION	AM615836				
KEYWORDS	AM615836.1	GI:7321721			
SOURCE	EST.				
ORGANISM	tomato.				
REFERENCE	Lycopersicon esculentum				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asterales; euasterids I; Solanales; Solanaceae; Solanum;				
TITLE	Lycopersicon.				
JOURNAL	1 (bases 1 to 574)				
COMMENT	van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang, .F., Hansen,T., Craeven,M.B., Bowman,C.L., Romling,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D. Generation of ESTs from tomato flower tissue, 0-3 mm buds Unpublished (1999) Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU 5 prime sequence.				

FEATURES	Location/Qualifiers
source	1. .574

```

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT0A17E20"
/clone_1lb="tomato flower buds 0-3 mm, Cornell University"
/tissue_type="flower"
/dev_stage="0-3mm buds"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tankeley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
BASE COUNT
ORIGIN
194 a          96 c          108 g          175 t          1 others

```

Query Match	1.5%	Score 103.4	DB 118	Length 574
Best Local Similarity	56.7%	Pred. No. 9.4e-08		
Matches 191	Conservative 0	Mismatches 146	Indels 0	Gaps 0

[illegible]

RESULT	15
LOCUS	AM615967
DEFINITION	AM615967 574 bp mRNA EST323333 tomato flower buds 0-3 mm, Cornell University
ACCESSION	Lycopersicon esculentum cDNA clone CTOA17E16 5', mRNA sequence.
VERSION	AM615967
KEYWORDS	AM615967.1 GI:7321720
ORGANISM	EST.
SOURCE	tomato.
REFERENCE	Lycopersicon esculentum Eukaryotes: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Asterales: euasterids I; Solanales; Solanaceae: Solanum; Lycopersicon. 1 (bases 1 to 574) van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang ,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D. Generation of ESTs from tomato flower tissue, 0-3 mm buds Unpublished (1999) Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU
JOURNAL	5 prime sequence.
COMMENT	

FEATURES	Location/Qualifiers
source	1. .574

```

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db.xref="taxon:4081"
/clone="crt0A17E16"
/clone_lib="tomato flower buds 0-3 mm, Cornell University"
/tissue_type="flower"
/dev_stage="0-3mm buds"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wk old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

Query Match	1.5%	Score	103.4	DB	118	length	574
Best Local Similarity	56.7%	Pred. No.	9.4e-08				
Matches	191	Conservative	0	Mismatches	146	Indels	0
						Gaps	0

OY	3496	tttctacaggtatggtlaagaalatalaataatcgaaacttglttggagaaccaacgaatcglalca	3555
Db	11	TGTTCTAGTATGGGAAGTGTTCAAGTCACATTATTTTTOCCCCAACACTGGTGACA	70
OY	3556	gcgtgatgctggaactaatagattcatatatcaaaccgaaggaaagctcttgyaatgagt	3615
Db	71	TGTACCAGAAGCCTTAATTACTTCATCATATTCAAAGGAAGATTAAGTTATTCAGGTAGT	130
OY	3616	tatctcgaagtagtaggtgagatcttcctggaaaatgctgagatcttgtcttglttgtagac	3675
Db	131	TATCCAAAGCCAAATTCATAGTATTACTTGGCAAAGTTTCATTGCCTTGCGCTGTGGCGAC	190
OY	3676	atgcatagagatagagaaglatctcgcttaacctcttaagtaacgcgaagctcttagaac	3735
Db	191	ACACATAAAGCGTTAGAGATGTTTCTATTATCACTATACGACCATTTAAGCTTAAGCT	250
OY	3736	attctacttaaagaatglttgsagacaacttgytltgtctctgatatcttgcgaacaaaaac	3795
Db	251	GAGTTTATTAATGATGTTTGAACAATTAGCACTTCACATTTCTCCAATCATGGAAGATATAA	310
OY	3796	tctattctctgctcaagcagccgaacaaagttt	3832
Db	311	CATCAAGTCANACTACGGGAAGGCAAGAAATTTTT	347

Mon Oct 1 11:34:40 2001

us-09-502-426-1.rst

Page 12

Search completed: September 29, 2001, 11:18:30  
Job time: 16099 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 06:50:21 : Search time 139.83 Seconds  
(without alignments)  
9325.431 Million cell updates/sec

Title: US-09-502-426-1

Perfect score: 6888

Sequence: 1 atgttggtattatattgttgccttagtcagccaagctt 6888

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92.8	1.3	4937	2	US-08-622-166A-3
2	87.6	1.3	1608	2	US-08-622-166A-1
3	77.4	1.1	19124	2	US-08-487-826B-13
4	68.6	1.0	837	4	US-08-998-416-288
5	67.4	1.0	19124	2	US-08-487-826B-13
6	65.8	1.0	615	4	US-08-998-416-186
7	65.4	0.9	636	4	US-08-998-416-1137
8	63.6	0.9	8920	2	US-08-446-855A-1
9	63.6	0.9	8920	2	US-09-150-741-1
10	62.6	0.9	4818	3	US-08-817-926-27
11	62.6	0.9	7218	1	US-08-232-463-14
12	61.8	0.9	5852	1	US-07-867-106-2
13	60.8	0.9	4818	3	US-08-817-926-27
14	60.6	0.9	837	4	US-08-998-416-288
15	60.4	0.9	8920	2	US-08-446-855A-1
16	60.4	0.9	8920	2	US-09-150-741-1
17	60.2	0.9	665	2	US-08-883-795A-36
18	59.2	0.9	51952	3	US-08-947-823-1
19	59	0.9	827	4	US-08-998-416-535
20	58.4	0.8	665	2	US-08-883-795A-36
21	57.8	0.8	1511	1	US-07-991-867B-8
22	57.8	0.8	1511	1	US-08-107-755A-8
23	57.8	0.8	1511	2	US-08-544-332-8
24	57.4	0.8	3095	6	5231168-1
25	56.4	0.8	2058	2	US-08-749-391-1
26	56.4	0.8	2058	3	US-09-390-200-1
27	56.4	0.8	4673	1	US-07-638-431-1

28	56.4	0.8	4673	5	PCT-US92-00018-1	Sequence 1, Appli
29	56.2	0.8	2960	3	US-08-913-842-3	Sequence 3, Appli
30	55.6	0.8	701	4	US-08-998-416-701	Sequence 701, App
31	55.6	0.8	5852	1	US-07-867-106-2	Sequence 2, Appli
32	55.2	0.8	1406	3	US-08-913-842-6	Sequence 8, Appli
33	55.2	0.8	1511	1	US-07-991-867B-8	Sequence 8, Appli
34	55.2	0.8	1511	1	US-08-107-755A-8	Sequence 8, Appli
35	55.2	0.8	1511	2	US-08-544-332-8	Sequence 8, Appli
36	55	0.8	1850	3	US-08-617-860B-32	Sequence 32, Appli
37	55	0.8	4098	2	US-08-605-106-4	Sequence 1, Appli
38	54.8	0.8	6768	1	US-08-107-755A-1	Sequence 1, Appli
39	54.8	0.8	8457	1	US-07-991-867B-1	Sequence 1, Appli
40	54.8	0.8	8457	2	US-08-544-332-1	Sequence 1, Appli
41	54.6	0.8	9048	3	US-08-973-273-4	Sequence 4, Appli
42	54.2	0.8	663	4	US-08-998-416-191	Sequence 191, App
43	53.8	0.8	6243	2	US-09-056-075-1	Sequence 1, Appli
44	53.6	0.8	2430	4	US-08-845-258-3	Sequence 3, Appli
45	53.6	0.8	2430	4	US-08-990-571-3	Sequence 3, Appli

#### ALIGNMENTS

RESULT 1  
US-08-622-166A-3  
: Sequence 3, Application US/08622166A  
: Patent No. 5952545  
: GENERAL INFORMATION:  
: APPLICANT: KONCZ, CSABA  
: APPLICANT: MATYUR, JADEEP  
: APPLICANT: SEKERES, MILOS  
: APPLICANT: ALTMANN, THOMAS  
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME  
: NUMBER OF SEQUENCES: 4  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
: STREET: P.O. Box 747  
: CITY: Falls Church  
: STATE: Virginia  
: COUNTRY: USA  
: ZIP: 22040-0747  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: OPERATING SYSTEM: IBM PC compatible  
: SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/622,166A  
: FILING DATE: 27-MAR-1996  
: CLASSIFICATION: 800  
: ATTORNEY/AGENT INFORMATION:  
: NAME: SVENSSON, LEONARD R.  
: REGISTRATION NUMBER: 30,330  
: REFERENCE/DOCKET NUMBER: 0147-0153P  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (703) 205-8000  
: TELEFAX: (703) 205-8050  
: TELEX: 248345  
: INFORMATION FOR SEQ ID NO: 3:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 4937 base pairs  
: TYPE: nucleic acid  
: STRANDEDNESS: double  
: TOPOLOGY: linear  
: MOLECULE TYPE: DNA (genomic)  
: HYPOTHETICAL: NO  
: ANTI-SENSE: NO  
: ORIGINAL SOURCE:  
: ORGANISM: Arabidopsis thaliana  
: STRAIN: cv. Columbia  
: IMMEDIATE SOURCE:  
: LIBRARY: lambda gt10

CLONE: C204  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(968..1483, 1680..1829, 1917..2165, 3903  
LOCATION: ..3989, 4084..4162, 4248..4354, 4446..4576, 4674  
LOCATION: ..4773)  
US-08-622-166A-3

Query Match 1.3%; Score 92.8; DB 2; Length 4937;  
Best Local Similarity 53.1%; Pred. No. 2.1e-09;  
Matches 225; Conservative 0; Mismatches 192; Indels 7; Gaps 1;

QY 3501 caggtatgtaagatatatagatgaactgttttgagaagaacagatcgtatcagctga 3560  
DB 1162 CCGGTACGGTTCGGTTTTCATGACGATCTTTTGGTGAACCCAGCATTTTCTCACCTGA 1221  
QY 3561 tgcgtacttaataagattcatattacaagaagaagcctcttgatgtatcttc 3620  
DB 1222 CCGGGAAGAACCGGTTTGTCTTTCAGAACGAGGAGCTTTTGAAGTGTCTTATCC 1281  
QY 3621 tagaagataagtgaggtatctcttgagaatggtcagatgctgtctgtgtgacatga 3680  
DB 1282 TGCTTCATTTTGAACCTTTTGGGGAACACCTCTGCTTCTTGAAGGTTCTTTGCA 1341  
QY 3681 tagagataagagaagatctcgtacttaactcttaagtcagcagcgtcttaagaacttct 3740  
DB 1342 TAAACCTATGACCTCTCTACCATGAGCTTTGCTAATTCCTCAATCAATCAAGCATCT 1401  
QY 3741 acttaagatglttgagaagacatacttctgttcttgatcttgcaacaacactat 3800  
DB 1402 CATGCTTGATTTGACCGGTTTACTCCGGTTTAACTTGATCTTG------TCTTCTC 1454  
QY 3801 ttctctgtcgaagacgaagcacaagaagtttattatcttattcttatttgtaatt 3860  
DB 1455 GTGTTCTGTCATGAGAAAGCAAAAGGTAAACAAAATCTTGCTTATCAAAAAC 1514  
QY 3861 ttctgttataagacttcttaagttcttaactttttttttaaattgaagattcagt 3920  
DB 1515 ATATATATTTATTTATTTATTCGGCTTCTCATCTATGTTTTTTTAAATAAATAAAT 1574  
QY 3921 ttaa 3924  
DB 1575 AAAA 1578

RESULT 2  
US-08-622-166A-1  
Sequence 1, Application US/08622166A  
Patent No. 5952545

GENERAL INFORMATION:  
APPLICANT: KONCZ, CSABA  
APPLICANT: MATHUR, JAIDEEP  
APPLICANT: SEKERES, MIKLOS  
APPLICANT: ALTMANN, THOMAS  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/622,166A  
FILING DATE: 27-MAR-1996

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: SVENSSON, LEONARD R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 0147-0153P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1608 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: A. thaliana

IMMEDIATE SOURCE:  
LIBRARY: lambda gt10  
CLONE: C204  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 48..1466  
US-08-622-166A-1

Query Match 1.3%; Score 87.6; DB 2; Length 1608;  
Best Local Similarity 56.6%; Pred. No. 1.6e-08;  
Matches 162; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 3501 caggtatgtaagatatatagatgaactgttttgagaagaacagatcgtatcagctga 3560  
DB 242 CCGGTACGGTTCGGTTTTCATGACGATCTTTTGGTGAACCCAGCATTTTCTCACCTGA 301  
QY 3561 tgcgtacttaataagattcatattacaagaagaagcctcttgatgtatcttc 3620  
DB 302 CCGGGAAGAACCGGTTTGTCTTTCAGAACGAGGAGCTTTTGAAGTGTCTTATCC 361  
QY 3621 tagaagataagtgaggtatctcttgagaatggtcagatgctgttctgtgtgagactga 3680  
DB 362 TGCTTCATTTTGAACCTTTTGGGGAACACCTCTGCTTCTTGAAGGTTCTTTGCA 421  
QY 3681 tagaagataagagaagatctcgtacttaactcttaagtcagcagcgtcttaagaacttct 3740  
DB 422 TAAAGTATGACCTCTCTACCATGAGCTTTGCTAATTCCTCAATCAATCAAGCATCT 481  
QY 3741 acttaagatglttgagaagacatacttctgttcttgatcttgg 3786  
DB 482 CATGCTTGATTTGACCGGTTTACTCCGGTTTAACTTGATCTTG 527

RESULT 3  
US-08-487-826B-13  
Sequence 13, Application US/08487826B  
Patent No. 5993827

GENERAL INFORMATION:  
APPLICANT: Sim, Kim L.  
APPLICANT: Childs, Chetan  
APPLICANT: Miller, Louis H.  
APPLICANT: Peterson, David S.  
APPLICANT: Su, Xin-zhaun  
APPLICANT: Wellens, Thomas E.  
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th floor  
CITY: Newport Beach  
STATE: California

```

COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121.001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

```

```

Query Match 1.1%; Score 77.4; DB 2; Length 19124;
Best Local Similarity 42.6%; Pred. No. 2.9e-06;
Matches 590; Conservative 0; Mismatches 781; Indels 14; Gaps 3;

```

```

QY 1013 ttataatcagattttacacgctcactaglttgagattactagcaataagcataagagacc 1072
DB 229 TTATATATCAATATATATATAGTTCCTTATTAATAATAATTAATATATACAAATAATAT 288
QY 1073 gttaagctatttatacaaaagttacaactgtaataagcttgaataatcccttgaaat 1132
DB 289 TTCTTGTTATTTTATTAATAATATATACATTAATTCCTTATTTTATTTATTAATTAATTCCTTTT 348
QY 1133 tggaaatccggttgcttgtaataatagatttagtgtaaaacaatagttatcaat 1192
DB 349 AATTTCTTAATTCCTTTTATGCAACAAAAACATAAGTAATTCATACATACACAAAA 408
QY 1193 agtgcacacataacataatcttcaacgaaacaaacaaacttaagagaagttacatat 1252
DB 409 AAAAAAATAAAAAAATAAAAAATTTATTAATTAATAAAAAATTAATAAGACATACGT 468
QY 1253 ccataatgggtatgctatccttcacgtacgtactagactagaagactaaagatgttat 1312
DB 469 TCACCTTATTAATTAATAATGATTTATTAACATTAATAACATATTGAGATTATTAATAATTA 528
QY 1313 gtgagtcgataaataagaaatcacacgctggttaataatlaatlaeggaacggttaacat 1372
DB 529 TTTTAACATTAAGAAAGATTAGAAATACATTTTATTTTATTTTATTTTGAATATTCATCAACAT 588
QY 1373 caatgcacataatcattctgtgtggtcaacaataaacaacaaacagaacagaacaa 1432
DB 589 ATATATATATATATATATCTTTTATTAATTAATAATAATAATTCCTTATTAATTAATTCCTT 648
QY 1433 cgattttcttgattccattcaatgataatgaatcagatccttttggtgtacagtt 1492
DB 649 TCTTTTATC-----ACATGTAATAATTAATAATAATAATTTTCGATTATTAATTCATATATTT 702
QY 1493 cgaagtcctcctacaacgctgtaacacatcgcaactlaaattgcttcttcaatgcac 1552
DB 703 ATGTGTTATTAATTAATTAATAGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 762
QY 1553 tttaacataattatgttagttggaatttaataagagcgaactgttaacattacatat 1612
DB 763 AATTAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 822

```

```

QY 1613 tatattagactagtagtatttccaaatacactcttgatgtttaaactatc 1672
DB 823 TTATTTCAAAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 882
QY 1673 ttgttcttcacaggtataaataatcaatcagcggtaaaaaagtttgcttatt 1732
DB 883 GCATATACATGCACATTTTAATTTTATTAATAAAAAATTTTATTTTATTAATTAATTAATTAACA 942
QY 1733 cgcagatgaagaagataaaccttaatgacttaatttttgaaatgtaaccccttact 1792
DB 943 ATTTCATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1002
QY 1793 catagataattaccgtagtttgggtgcataatgacagccttcaacgtgatgt 1852
DB 1003 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1062
QY 1853 caattttctcgaataataatgaagaatcactacatcactagaagaacagc 1912
DB 1063 CTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1122
QY 1913 tgaattacacatttaatttaagaacaaatlttgaaatgtaataattctaacat 1972
DB 1123 TTAATAAATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1182
QY 1973 attataaataatgtagtactaagatgttccctatgcttcttaaatattttttata 2032
DB 1183 TAAGATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1237
QY 2033 tttagttataacatcactatgaacaaataatagttggtaattcacaatccat 2092
DB 1238 TTATATATAGTATATAGTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1294
QY 2093 atttttgaacatcaaatlaataatlttagtcaataacaaatgcatagaagttccaa 2152
DB 1295 TTTGTATTAACATTAACAAAGACTTAAGAAACTATACATCTGATCTATTAATTAATTAATTAAT 1354
QY 2153 aaaaaatttgtaacagaaacttcaaatlttttttttttttttttttttttttttttttttttttt 2212
DB 1355 ATTAATATCTTTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1414
QY 2213 atagaacactatgttggtggaagtagtaataacataacataacataacataacataacataacata 2272
DB 1415 TAATATATTTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1474
QY 2273 attataaagccttaacgctcaagtaatgttacttagttaggttaataatgcat 2332
DB 1475 AAAACCTTCAAAACATTTTTCATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1534
QY 2333 ggtgcagatcagaattgggaacacaaatgaaacggaatlaaataatlaacttaataa 2392
DB 1535 AGAGAGAAACGTAGAACATTAACCAAAAAAATTAAGAACAAAAAATTAATTAACAAAAAATAA 1594
QY 2393 ataaa 2397
DB 1595 TAAAA 1599

```

```

RESULT 4
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
GENERAL INFORMATION:
APPLICANT: Philippsen, Peter
APPLICANT: Pohlmann, Rainer
APPLICANT: Steiner, Sabine
APPLICANT: Mohr, Christine
APPLICANT: Wendland, Jurgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152

```









Db 384 TTATTTATTTTAAATTAATCATTTTAAATTTAAATATATGTTGATTAATATTAATTTAAAT 443

QY 2260 caaatlttaaaaaataatataaagcctatacgcgcctcaagaatagtatctagtagtgta 2319

Db 444 CTTTTATTAGAAATTTATTTAAATAATTAATTTTAATCTTAAATTCCTATTTATTAATTTT 503

QY 2320 attaatatgcatggtggcgattcaagaattggggcaacaatgaaaaaggaaatlaaataatt 2379

Db 504 ATATTTATTTAAATAATTAATTAATTCATTTTATTTATTTATTTAAATTAATTAATTTAT 563

QY 2380 aacttaaaataaataaaatttgcgtaaat 2410

Db 564 TAATTAATATTTATCATCTTTTAAATTAAT 594

RESULT 8  
 US-08-446-855A-1  
 : Sequence 1, Application US/08446855A  
 Patent No. 5849573  
 GENERAL INFORMATION:  
 APPLICANT: Stewart, Thomas S  
 - APPLICANT: Flores, Maria V  
 APPLICANT: O'Sullivan, William J  
 TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl  
 TITLE OF INVENTION: phosphate synthetase II  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Nixon & Vanderhye PC  
 STREET: 1100 No. 5849573th Glebe Road, 8th Floor  
 City: Arlington  
 STATE: Virginia  
 COUNTRY: USA  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentln Release #1.24  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/446,855A  
 FILING DATE: 06-Jul-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Milchard, Leonard C  
 REGISTRATION NUMBER: 29.009  
 REFERENCE/DOCKET NUMBER: 47-80  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-816-4000  
 TELEFAX: 703-816-4100  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 8920 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: genomic  
 US-08-446-855A-1

	Query Match	Similarity	Best Local	Matches	Conservative	Score	DB 2	Length	Gaps
		46.8%		278;	0;	63.6;		8920;	
						Pred. No.	0.0011;		
						Mismatches	304;		
						Indels	12;		
						Gaps	2;		
Oy	1874	aattggaattcaatgcgtactctacatagaaagaacagctgagtattacatttaatta	1933						
Db	152	AAAAAAAAATATAAAAAAATAAAAAAAGATTAAAAAATTTTATATTTCATGTGATGTAT	211						
Oy	1934	aagacaataatttgaataatgata-attctaacaatatataaataatgatgcgt	1992						
Ddb	212	ATATATATCATTTTCATATTATTAATACATGTGAATAAAAAACCCTTTTTTTTTTTTTTTTCT	271						
Oy	1993	ataatgatttccttatgctctttaaacattttttlllataattagttacaataacattat	2052						

Db	272	TTATATTTATTAACAAATACCTTTTAAGTTATTTATATATATATATATATATATATATA	331
QY	2053	gaaccaataaagtcgtggaattcaaatatctccatlaaattcttggaaatcacaat	2113
Db	332	TATATATATATATATGTTTGTGTTCAATTTGTTATTAATAATTAAGTGAATAATAACT	391
QY	2113	tatbaatatttaglcaataaacaatgcataagaaagtcacaaaaaaatttgttaacgaa	2173
Db	392	TATTAATATATTTCCAAATTAATATGAATACAAATTTAATATTTTCATGTGCACATTA	451
QY	2173	acttccaaatttttttttttattgga-----acaagaataacagatgagaac	2223
Db	452	ATATAGTTTACACTCTTATTAATAAACCATCTATATATATACCAATATATATATAC	511
QY	2222	tatttctgttggaatggaagtcgataataacattgaagcaatttcaaaaaattatata	2283
Db	512	TCCCAATATGTGGTCCATATATTTTATTAATATATATATATATATATATATATAT	571
QY	2282	gcctacgcgcgcacaagtcgtctatccttagtgtagtgaataaagtcgagtcgatt	2341
Db	572	ATTTATTTTCTTGAAGTTATTAATAATAGTAATCTACTAATTTAAAAA	631
QY	2342	cagaattcggaacacaaatgaaaacggaattaaataatlaacttlaaatlaaataaatt	2403
Db	632	AAAAAAAAAAAAAAAAAAAAAAAAATTTACATGAAAAATGAACCTTGTAATGTAA	691
QY	2402	tgaagtaagtcgtcttctgactatcgaggggcacaaaagaacatgcgcaaaag	2455
Db	692	TTTATTAATTTTAAACATTAATATATATATGCTTAAAAA	745

```

RESULT      9
US-09-150-741-1
; Sequence 1, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; Patent No. 6183996
; TITLE OF INVENTION: Synthetase II
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; CURRENT FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PL6380
; EARLIER FILING DATE: 1992-12-16
; EARLIER APPLICATION NUMBER: A093/00617
; EARLIER FILING DATE: 1993-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8920
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; US-09-150-741-1

```

	Query Match	Best Local Similarity	0.9%; 46.8%;	Score 63.6; Pred. NO. 0.0011;	DB 4;	Length 8920;
	Matches	278;	Conservative	0;	Mismatches 304;	Indels 12; Gaps 2.
QY	1874	aattggaatcgaatgcctactcaatagaagaacagctgagttatcattttaatta	1933			
DB	152	aaaaaaaaataaataaaaaataaaagaatataaaatatttatattgatatgcagat	211			
OY	1934	aagacaaaattttgaaaaatgltata-attctacaacatatattaataatgagcct	1992			
DB	212	ataatagatattcatattcaattaacaacagataaaaacattttttttttttctct	271			
OY	1993	ataagattttccctatgctttcaaaaataatttttttttatattgattgattaaaataacatcat	2052			
DB	272	ttataatttataacatacatcatttaagttattttatataataataataataataata	331			





Db 5607 AAAAAAAAAAATGTAATTTGCAATTAATAAATGTAAGCGTTTTTTT 5548  
QY 2007 atgtcttaaaatatttttttata-----ttaagtaaaatacatgaacca 2059  
Db 5547 AAAATTATGTCATGATTTTATTTTAAATGACGAGATTAATAATCTTAAACAA 5488  
QY 2060 taatagtcgtggaattcaaatatctccatataatttttgaatctacaattataat 2119  
Db 5487 TAAACATATTTGATTTTTTTTTTTTTTTTTTTTTTAAATCAATTAATTAATA 5428  
QY 2120 attagtaacaatacatagagaagttccaaaaaatttgtaacagaactcca 2179  
Db 5427 TAATATCTATATATCTTGATGAACTTCATTTTATTAATTAATTTAATTTTAA 5368  
QY 2180 aattttttttt 2192  
Db 5367 AATATGATCT 5355

RESULT 13  
US-08-817-926-27/c  
Sequence 27, Application US/08817926  
Patent No. 6001590  
GENERAL INFORMATION:  
APPLICANT: Komeda, Toshinori  
APPLICANT: Suda, Hisako  
APPLICANT: Tamai, Yukio  
APPLICANT: Imamatsu, Akihiro  
APPLICANT: Kato, No. 6001590uo  
APPLICANT: Sakai, Yasuyoshi  
TITLE OF INVENTION: PROMOTER/TERMINATOR FOR CANDIDA BOIDINI  
TITLE OF INVENTION: FORMATE DEHYDROGENASE GENE  
NUMBER OF SEQUENCES: 51  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/817,926  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/02597  
FILING DATE: 12-SEP-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 234133/1995  
FILING DATE: 12-SEP-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 42536/1996  
FILING DATE: 29-FEB-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 29,768  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4818 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Candida boidini  
US-08-817-926-27

Query Match 0.98; Score 60.8; DB 3; Length 4818;  
Best Local Similarity 46.38; Pred. No. 0.0033;  
Matches 275; Conservative 0; Mismatches 312; Indels 7; Gaps 2;

QY 54 ttcggtattcctgcgtccatgattgagtttgcttcacatcgttcgaagaataat 113  
Db 4718 TTTTATTTATTTCTTTTATTTTCTTATATATGTCATTTTGGCTTATTAACAATA 4659  
QY 114 aaatcaataatcattcaataataattcaagaatcaatcattcattcattgatt 173  
Db 4658 AAATTTAGAGTCCTTTTAAAGTTTTCGATTAATTAATTAATTAATTAATAGGT 4599  
QY 174 atacaataatcattcattcattgattgattgattgattgattgattgattgatt 233  
Db 4598 AAATAGAGTAAATAAATAAATGATACCAATTTAAACAAATTTAATTAATTAATA 4539  
QY 234 a-aggaataatcgtatccattcattgattgattgattgattgattgattgattgatt 292  
Db 4538 TTTTAAATTTTAAACGTTTGTTCAAAAAATTTAGAAATTTAATTAATTAATTA 4479  
QY 293 ttatattatattgattcaatacaaaagattgattgattgattgattgattgattgatt 352  
Db 4478 CAACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4419  
QY 353 gaccccaaaaaaataaaaaaatacaaaacaaaccccccccgatattgattgatt 412  
Db 4418 TGATACAAAAAATAAATAAACAAGAAATTAACAACAAACAAACAAATTTTCTTTT 4359  
QY 413 tctgattgattgattgattgattgattgattgattgattgattgattgattgatt 466  
Db 4358 TTTTATTTACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4239  
QY 467 gaagatttctcaataataaattcgaattcattcattcattcattcattcattcatt 526  
Db 4298 AAAGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4239  
QY 527 gattgattcattcattcattcattcattcattcattcattcattcattcattcatt 586  
Db 4238 TACTCTTTTAAACGTTTGAACATAATTAATTAATTAATTAATTAATTAATTAATTC 4179  
QY 587 gcaattattcaagaagattgattgattgattgattgattgattgattgattgatt 640  
Db 4178 TTCAATTTCAATTTTAAATTTGAATAAACAAGAAATTAATAAACAAGAAATTA 4125

RESULT 14  
US-08-998-416-288/c  
Sequence 288, Application US/08998416  
Patent No. 6239264  
GENERAL INFORMATION:  
APPLICANT: Philippsen, Peter  
APPLICANT: Pohlmann, Rainer  
APPLICANT: Steiner, Sabine  
APPLICANT: Mohr, Christine  
APPLICANT: Wendland, Jurgen  
APPLICANT: Knechtie, Philipp  
APPLICANT: Reibischung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPYII  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NO. 6239264ch Carolina  
COUNTRY: USA  
ZIP: 27709

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/N/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 288:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1241RP
US-08-998-416-288
```

Query Match 0.9%; Score 60.6; DB 4; Length 837;  
Best Local Similarity 49.1%; Pred. No. 0.0023;  
Matches 191; Conservative 0; Mismatches 194; Indels 4; Gaps 1;

```
QY 1939 aaaaatttggaaaatgataatttctacaatatatttaaatatgatgacctataatg 1998
DB 608 AATAATATTTTAAATTAATTAATTAATGAATTAATTAATTAATTAATTAATTAAT 549
QY 1999 tatttcctagtccttaaaatattttttattattagtttaataatcatatgacaa 2058
DB 548 AATAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 489
QY 2059 ataatagttgatgaatcaaatatccatcaatatttttgaatctacaataatttaa 2118
DB 488 GAAATTAAGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 429
QY 2119 tatttagtaatacaatgcatagaaagttccaaaaaaatttgttaacagaacttcc 2178
DB 428 TCAACATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 369
QY 2179 aaaaatttctttaaagaaacagaatacagatagaactatttgggtggaatg 2238
DB 368 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 313
QY 2239 gaagtagtaataacatgaacaaatttaaaaaattataaagcctacaagcgtcaaa 2298
DB 312 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 253
QY 2299 gtaagttaactagtggttaattataa 2327
DB 252 AATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 224
```

RESULT 15  
US-08-446-855A-1/C  
Sequence 1, Application US/08446855A  
Patent No. 5849573  
GENERAL INFORMATION:  
APPLICANT: Stewart, Thomas S  
APPLICANT: Flores, Maria V  
APPLICANT: O'Sullivan, William J  
TITLE OF INVENTION: Nucleotide sequence encoding cardamyl

```
TITLE OF INVENTION: phosphate synthetase II
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: Nixon & Vanderhye PC
STREET: 1100 NO. 5849573th Glebe Road, 8th Floor
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,855A
FILING DATE: 06-Jul-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 47-80
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic
US-08-446-855A-1
```

Query Match 0.9%; Score 60.4; DB 2; Length 8920;  
Best Local Similarity 48.8%; Pred. No. 0.0047;  
Matches 163; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

```
QY 1919 ttacatttaaatgaagcaaaatttggaaaatgataatttctacaatatattt 1978
DB 8751 TTTTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8692
QY 1979 aaaaatgagcctaataatgataatttccatagttcttaaaatattttttattattg 2038
DB 8691 AAACCATTTTGGTTATACATATGATGAATTAATTAATTAATTAATTAATTAATTA 8632
QY 2039 tataaatcaatatagaacaaatagttggtgaattcaaatatatttccattatattt 2098
DB 8631 TATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8572
QY 2099 tgaattctacaattatataatttagtcaatcaaatgcatagaaagttccaaaaaaa 2158
DB 8571 TTAACGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8512
QY 2159 ttctgttaacagaacttccaaatttttttttaagaaacagaataacagatagaa 2218
DB 8511 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8452
QY 2219 aactatttgcgtggaatgaagtagtaata 2252
DB 8451 AATATCAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8418
```

Search completed: September 29, 2001, 11:27:42  
Job time: 16641 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 07:23:36 ; Search time 378.88 seconds  
(without alignments)  
11415.168 Million cell updates/sec

Title: US-09-502-426-1

Perfect score: 6888  
Sequence: 1 atcgtggtattatattgttgccttagtcagccaaagcct 6888

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: N\_Geneseq\_0601:\*  
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:\*  
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:\*  
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:\*  
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:\*  
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:\*  
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:\*  
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:\*  
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:\*  
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:\*  
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:\*  
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:\*  
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:\*  
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:\*  
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:\*  
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:\*  
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:\*  
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:\*  
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:\*  
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:\*  
21: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:\*  
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:\*  
23: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6888	100.0	6888	21	AAA59599
2	251	3.6	936	22	AAF58252
3	251	3.6	936	22	AAF58254
4	251	3.6	936	22	AAF58257
5	251	3.6	936	22	AAF58259
6	251	3.6	936	22	AAF58262
7	251	3.6	936	22	AAF58255
8	250.6	3.6	936	22	AAF58252
9	250.6	3.6	936	22	AAF58254
10	250.6	3.6	936	22	AAF58257
11	250.6	3.6	936	22	AAF58259

#### ALIGNMENTS

12	250.6	3.6	936	22	AAF58262	Oligonucleotide D2
13	250.6	3.6	936	22	AAF58255	Oligonucleotide D1
14	92.8	1.3	4937	18	AAAT85307	Arabidopsis thalia
15	88.2	1.3	244	22	AAF58238	Oligonucleotide D1
16	87.6	1.3	1608	18	AAAT85306	Arabidopsis thalia
17	87.6	1.3	1646	21	AAAT85307	Arabidopsis thalia
18	87.6	1.3	1649	21	AAAT85306	Arabidopsis thalia
19	85.4	1.2	244	22	AAF58238	Oligonucleotide D1
20	79.4	1.2	8310	20	AAAT29911	CDNA encoding a SC
21	77.4	1.1	19124	18	AAAT72882	Plasmodium var-7 g
22	77.4	1.1	19124	21	AAAT72882	Plasmodium var-7 p
23	74	1.1	2104	13	AAAT25273	Sequence encoding
24	73	1.1	1864	8	AAAT71405	Sequence of ANS-1
25	70.8	1.0	8310	20	AAAT29911	CDNA encoding a SC
26	69.6	1.0	4590	7	AAAT6472	Sequence encoding
27	68.8	1.0	875	21	AAAT01920	Human colon cancer
28	68.6	1.0	1864	8	AAAT71405	Sequence of ANS-1
29	68.2	1.0	2104	13	AAAT25273	Sequence encoding
30	67.4	1.0	5760	6	AAAT50530	Sequence encoding
31	67.4	1.0	19124	18	AAAT72882	Plasmodium var-7 g
32	67.4	1.0	19124	21	AAAT72882	Plasmodium var-7 p
33	67.2	1.0	875	21	AAAT01920	Human colon cancer
34	66.6	1.0	1398	21	AAAT8714	Arabidopsis thalia
35	66	1.0	2486	21	AAAT97037	Nucleotide sequenc
36	65.8	1.0	2486	21	AAAT97037	Nucleotide sequenc
37	65.4	0.9	2503	15	AAAT53480	PNPX30 xylanase cd
38	65.4	0.9	20674	21	AAAT58017	Arachidonic acid m
39	65.2	0.9	6033	21	AAAT0152	Plasmodium falcipa
40	65	0.9	2418	13	AAAT27886	P.falciparum GBP3
41	64.4	0.9	5409	21	AAAT0151	Plasmodium falcipa
42	64.2	0.9	1431	21	AAAT37082	DNA sequence encod
43	64.2	0.9	1671	13	AAAT24134	50 KD subunit of S
44	63.6	0.9	8920	15	AAAT62924	Carbamoyl-phosphat
45	63	0.9	20674	21	AAAT58017	Arachidonic acid m

#### RESULT 1

AAAT59599	standard; DNA; 6888 BP.
ID	AAAT59599
XX	AAAT59599;
AC	14-NOV-2000 (first entry)
XX	
DT	DNA encoding a cytochrome P450 enzyme designated DMF4.
DE	DMF4; cytochrome P450 enzyme; brassinosteroid; 22alpha-hydroxylation;
XX	plant phenotype; cell elongation; ss.
KW	Arabidopsis sp.
XX	
OS	Location/Qualifiers
XX	1..3203
FT	Key
FT	promoter
FT	/*tag= a
FT	3060..3125
FT	/*tag= b
FT	CDS
FT	3203..6110
FT	/*tag= c
FT	/product= "DMF4"
FT	/note= "contains introns"
FT	3203..3423
FT	/*tag= d
FT	3424..3503
FT	/*tag= e
FT	3504..3828
FT	/*tag= f
FT	3829..3913
FT	/*tag= g
FT	3914..4066
FT	/*tag= h
FT	exon





QY	1261	gggtatgctataccttccacgtaatgctatactagagactaaagaatgtagtctgctgc	1320
DB	1261	gggtatgctataccttccacgtaatgctatactagagactaaagaatgtagtctgctgc	1320
QY	1321	gataaatgaaatccacacgctgggtaataatataaggacgtatgttccgactcgtcaa	1380
DB	1321	gataaatgaaatccacacgctgggtaataatataaggacgtatgttccgactcgtcaa	1380
QY	1381	ataatcattctgtgtgtgtccaaataaataaacaagaataaagaataaagaatcttc	1440
DB	1381	ataatcattctgtgtgtgtccaaataaataaacaagaataaagaataaagaatcttc	1440
QY	1441	cttgatccatcccaatgataatgataatgataatgataatgataatgataatgataat	1500
DB	1441	cttgatccatcccaatgataatgataatgataatgataatgataatgataatgataat	1500
QY	1501	tcataaagcgtgtgaacacacgcgtcaactataatgtcttcttcaatgacatcttaac	1560
DB	1501	tcataaagcgtgtgaacacacgcgtcaactataatgtcttcttcaatgacatcttaac	1560
QY	1561	attatgt	1620
DB	1561	attatgt	1620
QY	1621	atactagatgt	1680
DB	1621	atactagatgt	1680
QY	1681	tcctacggtataaataatcaatcaatcagagtaaaaaagtttgccttatttcgcgacgc	1740
DB	1681	tcctacggtataaataatcaatcaatcagagtaaaaaagtttgccttatttcgcgacgc	1740
QY	1741	atgagagtaaaccttaagcttaatttttttgaagaatgtaaaccttcttcttcttcttct	1800
DB	1741	atgagagtaaaccttaagcttaatttttttgaagaatgtaaaccttcttcttcttcttct	1800
QY	1801	aatacgcgtatgt	1860
DB	1801	aatacgcgtatgt	1860
QY	1861	tcctgcaaatataaataatgataatgataatgataatgataatgataatgataatgata	1920
DB	1861	tcctgcaaatataaataatgataatgataatgataatgataatgataatgataatgata	1920
QY	1921	acatttaatttaaaagcaaaatgttttttgaagaatgtaaaccttcttcttcttcttct	1980
DB	1921	acatttaatttaaaagcaaaatgttttttgaagaatgtaaaccttcttcttcttcttct	1980
QY	1981	aataatgctataatgataatgataatgataatgataatgataatgataatgataatgata	2040
DB	1981	aataatgctataatgataatgataatgataatgataatgataatgataatgataatgata	2040
QY	2041	taataaactatgataatgataatgataatgataatgataatgataatgataatgataatg	2100
DB	2041	taataaactatgataatgataatgataatgataatgataatgataatgataatgataatg	2100
QY	2101	aaatcaaaatataatgataatgataatgataatgataatgataatgataatgataatg	2160
DB	2101	aaatcaaaatataatgataatgataatgataatgataatgataatgataatgataatg	2160
QY	2161	ttgttaacagaacttccaaatgttttttttgaagaatgtaaaccttcttcttcttcttct	2220
DB	2161	ttgttaacagaacttccaaatgttttttttgaagaatgtaaaccttcttcttcttcttct	2220
QY	2221	ctattttgt	2280
DB	2221	ctattttgt	2280
QY	2281	agccatacgcgtcaaaagtatgtatctagttaggtgtaataatgataatgataatgata	2340
DB	2281	agccatacgcgtcaaaagtatgtatctagttaggtgtaataatgataatgataatgata	2340
QY	2341	tcagaattgggacaacatgaaaaaggaataaataatgataatgataatgataatgata	2400
DB	2341	tcagaattgggacaacatgaaaaaggaataaataatgataatgataatgataatgata	2400
QY	2401	ttgagtaaatgt	2460
DB	2401	ttgagtaaatgt	2460
QY	2461	gggttgcgt	2520
DB	2461	gggttgcgt	2520
QY	2521	taggggtccttcgataatgataatgataatgataatgataatgataatgataatgata	2580
DB	2521	taggggtccttcgataatgataatgataatgataatgataatgataatgataatgata	2580
QY	2581	tatcctaaggtaatttgggaatcccaatttaaacgagatgagacgttacccgacttcc	2640
DB	2581	tatcctaaggtaatttgggaatcccaatttaaacgagatgagacgttacccgacttcc	2640
QY	2641	tggaattctgt	2700
DB	2641	tggaattctgt	2700
QY	2701	aactcaaaccttgatcagataaataatgataatgataatgataatgataatgataatg	2760
DB	2701	aactcaaaccttgatcagataaataatgataatgataatgataatgataatgataatg	2760
QY	2761	atctaatgacttttttttcttccacacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2820
DB	2761	atctaatgacttttttttcttccacacggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	2820
QY	2821	acaattgtatgataatgataatgataatgataatgataatgataatgataatgataatg	2880
DB	2821	acaattgtatgataatgataatgataatgataatgataatgataatgataatgataatg	2880
QY	2881	ctatttcgacatcgcagatcttctgcgaacttctgtgtgtgtgtgtgtgtgtgtgtgt	2940
DB	2881	ctatttcgacatcgcagatcttctgcgaacttctgtgtgtgtgtgtgtgtgtgtgtgt	2940
QY	2941	aagtaataaaggaagcaatacgaaggaaggaaggaaggaaggaaggaaggaaggaag	3000
DB	2941	aagtaataaaggaagcaatacgaaggaaggaaggaaggaaggaaggaaggaaggaag	3000
QY	3001	tgaaatcattgt	3060
DB	3001	tgaaatcattgt	3060
QY	3061	ataaacaattataatcttgcgaacttctttaaaccctcttacaacagagatcccttct	3120
DB	3061	ataaacaattataatcttgcgaacttctttaaaccctcttacaacagagatcccttct	3120
QY	3121	tcagtaagaagcccgatcttcccaatcttaaaagacacatgagaagaagtgtgtgtgt	3180
DB	3121	tcagtaagaagcccgatcttcccaatcttaaaagacacatgagaagaagtgtgtgtgt	3180
QY	3181	gagagagagagaaactagatcctcaatgttcgaaacagagacatccttcccttctct	3240
DB	3181	gagagagagagaaactagatcctcaatgttcgaaacagagacatccttcccttctct	3240
QY	3241	tccttcccatcgtcttgcgtcttcttcttcttcttcttcttcttcttcttcttcttct	3300
DB	3241	tccttcccatcgtcttgcgtcttcttcttcttcttcttcttcttcttcttcttcttct	3300
QY	3301	aaaaacacagatccaatcctcccggtgtaaatccggttgcgaattcttcttcttcttct	3360
DB	3301	aaaaacacagatccaatcctcccggtgtaaatccggttgcgaattcttcttcttcttct	3360
QY	3361	cgttatctttaaaccgttacaacgcgaacacactcgtgtacttcaagcaacacatgtctc	3420
DB	3361	cgttatctttaaaccgttacaacgcgaacacactcgtgtacttcaagcaacacatgtctc	3420
QY	3421	caagtaaacacaacatcttccaaaactcaaaaaataaataatccttcttcttcttct	3480
DB	3421	caagtaaacacaacatcttccaaaactcaaaaaataaataatccttcttcttcttct	3480

Db 3421 caagtaacaacaacatctccaaaaactcaaaaaataaactcgtcttcttgaattc 3480  
Qy 3481 gactaatgtgttatttattcaagtaatgtaagataatagatcgaactgttttgagaa 3540  
Db 3481 gactaatgtgttatttattcaagtaatgtaagataatagatcgaactgttttgagaa 3540  
Qy 3541 ccaacagtcgtatcagctgtgtctgtgaacttaataagatcataltaacaaaagagagag 3600  
Db 3541 ccaacagtcgtatcagctgtgtctgtgaacttaataagatcataltaacaaaagagagag 3600  
Qy 3601 cctcttaagatgtatttccctagaagatagtgaggatctcttgagaaatgctcgaatcct 3660  
Db 3601 cctcttaagatgtatttccctagaagatagtgaggatctcttgagaaatgctcgaatcct 3660  
Qy 3661 gtctctgtgtgtagacatgcaatagatagatagagatcctcgtcttaacttcttaagtac 3720  
Db 3661 gtctctgtgtgtagacatgcaatagatagatagagatcctcgtcttaacttcttaagtac 3720  
Qy 3721 gcaacgtcttagaactatctcaacttaagaatgttgagagaaatactgtttcttgat 3780  
Db 3721 gcaacgtcttagaactatctcaacttaagaatgttgagagaaatactgtttcttgat 3780  
Qy 3781 tcttggaacaaacatctatttctcgtcgaagacgagccaaaagttttattttt 3840  
Db 3781 tcttggaacaaacatctatttctcgtcgaagacgagccaaaagttttattttt 3840  
Qy 3841 atcttatttctgtaaatatttctgttatagaactcttagagtttcttaactttttt 3900  
Db 3841 atcttatttctgtaaatatttctgttatagaactcttagagtttcttaactttttt 3900  
Qy 3901 tttaattgaaacagtttaacgttttaacttaaatgtgcgaagacataatagatagatcctg 3960  
Db 3901 tttaattgaaacagtttaacgttttaacttaaatgtgcgaagacataatagatagatcctg 3960  
Qy 3961 agagaagaagaacgagaaatcaataagaagaagatgttaacttcaatgaagaagttgctc 4020  
Db 3961 agagaagaagaacgagaaatcaataagaagaagatgttaacttcaatgaagaagttgctc 4020  
Qy 4021 tgcctccttaaatctacacaggaactgtctataaaagctcttcgaatattttt 4080  
Db 4021 tgcctccttaaatctacacaggaactgtctataaaagctcttcgaatattttt 4080  
Qy 4081 ttttgctgtaagatcacaactctcatatagtttttaattttatttttttggttaac 4140  
Db 4081 ttttgctgtaagatcacaactctcatatagtttttaattttatttttttggttaac 4140  
Qy 4141 aaaaatcctaaatgtgtgtgtagtgcagagcaagacatctgaagttcatctgagagaa 4200  
Db 4141 aaaaatcctaaatgtgtgtgtagtgcagagcaagacatctgaagttcatctgagagaa 4200  
Qy 4201 atggaagaagaagaatctgatacaagaagaatcaagaagaagaagaagtgaanaa 4260  
Db 4201 atggaagaagaagaatctgatacaagaagaatcaagaagaagaagaagtgaanaa 4260  
Qy 4261 gaagatgaacagagatgataagaagatcatgtttaggaacaaaagaacagcgatgat 4320  
Db 4261 gaagatgaacagagatgataagaagatcatgtttaggaacaaaagaacagcgatgat 4320  
Qy 4321 cttttaggaatgggtcttgaanaacttcgaatttcaacgagagcaaatctcgcattc 4380  
Db 4321 cttttaggaatgggtcttgaanaacttcgaatttcaacgagagcaaatctcgcattc 4380  
Qy 4381 ctttaggttatttctgcggaacatgagacttctcttgaacattgtctcgcattcttc 4440  
Db 4381 ctttaggttatttctgcggaacatgagacttctcttgaacattgtctcgcattcttc 4440  
Qy 4441 ttcttggaacgttgcccttaagacgttgaagaacttagaggttaagataatataacagcac 4500  
Db 4441 ttcttggaacgttgcccttaagacgttgaagaacttagaggttaagataatataacagcac 4500  
Qy 4501 aagttaattactacaacatgtttagatcatataaagtattatagatattctatag 4560  
Db 4501 aagttaattactacaacatgtttagatcatataaagtattatagatattctatag 4560

Qy 4561 aatatacgtatgaanaaagtatgatatatttaatttgcataatttcaatgattatgattta 4620  
Db 4561 aatatacgtatgaanaaagtatgatatatttaatttgcataatttcaatgattatgattta 4620  
Qy 4621 tacttttgaagaagaacatcttgagatcgcgagggccaaagaagaactagagatcag 4680  
Db 4621 tacttttgaagaagaacatcttgagatcgcgagggccaaagaagaactagagatcag 4680  
Qy 4681 aatataatggatgtatgaagaanaatggaacttcaactgaatgtgatactatcatt 4740  
Db 4681 aatataatggatgtatgaagaanaatggaacttcaactgaatgtgatactatcatt 4740  
Qy 4741 ctcattattatctatgltcaatagattatgaatgaanaacaaatattgatttttt 4800  
Db 4741 ctcattattatctatgltcaatagattatgaatgaanaacaaatattgatttttt 4800  
Qy 4801 ttggtgtgtgtgaaggtataaaatgaacacttcgatttgggaaatgtagttgatttg 4860  
Db 4801 ttggtgtgtgtgaaggtataaaatgaacacttcgatttgggaaatgtagttgatttg 4860  
Qy 4861 catcgcaaacgactcacaagaatgtctcgtaacaagaatgaacacttcaatcaaatttta 4920  
Db 4861 catcgcaaacgactcacaagaatgtctcgtaacaagaatgaacacttcaatcaaatttta 4920  
Qy 4921 aataatgaatccggaatattgaatacttatttgaatgaanaaataataaaattatcat 4980  
Db 4921 aataatgaatccggaatattgaatacttatttgaatgaanaaataataaaattatcat 4980  
Qy 4981 ttcttattgtttggaanaaagaatagatatacctctagttgggttgaagaagtgttacgggtgat 5040  
Db 4981 ttcttattgtttggaanaaagaatagatatacctctagttgggttgaagaagtgttacgggtgat 5040  
Qy 5041 ctacgccgtatcttgatatactctgtatagaaccaacccaactcttcaatccttgag 5100  
Db 5041 ctacgccgtatcttgatatactctgtatagaaccaacccaactcttcaatccttgag 5100  
Qy 5101 atggaagaagaatgaanaaagtctctcgttgaactacccaatagttgattttt 5160  
Db 5101 atggaagaagaatgaanaaagtctctcgttgaactacccaatagttgattttt 5160  
Qy 5161 ttatcatctgcatgtaataagatcgtccctacgtgatttacctctctataataatcag 5220  
Db 5161 ttatcatctgcatgtaataagatcgtccctacgtgatttacctctctataataatcag 5220  
Qy 5221 agaaacttaaatctggaacaaacttttattatattatattatattcttcaagttagatcttaaa 5280  
Db 5221 agaaacttaaatctggaacaaacttttattatattatattatattcttcaagttagatcttaaa 5280  
Qy 5281 cgtcatggaatcatctgatacaatgaagaacatataaattaaagaagaagaagagac 5340  
Db 5281 cgtcatggaatcatctgatacaatgaagaacatataaattaaagaagaagaagagac 5340  
Qy 5341 gtgtgtgtgaagtgttagcgtgaagaacatatttagtaggaatgtagtatttcaagag 5400  
Db 5341 gtgtgtgtgaagtgttagcgtgaagaacatatttagtaggaatgtagtatttcaagag 5400  
Qy 5401 taggaactgacctaaattcttataatcgtgacggttgatttatttagtcaagcgtt 5460  
Db 5401 taggaactgacctaaattcttataatcgtgacggttgatttatttagtcaagcgtt 5460  
Qy 5461 gatacaatttaagaatcttgaagaagatgtgataccaagcgtttgacggaatcccaacgccc 5520  
Db 5461 gatacaatttaagaatcttgaagaagatgtgataccaagcgtttgacggaatcccaacgccc 5520  
Qy 5521 atgactgtgacgttgcccttaagatctgcacagcttgacaggtgttttaacttagatttaa 5580  
Db 5521 atgactgtgacgttgcccttaagatctgcacagcttgacaggtgttttaacttagatttaa 5580  
Qy 5581 cttcatattataacaacatcatcaaatcggcggaatttagaganaaatalactatagattta 5640  
Db 5581 cttcatattataacaacatcatcaaatcggcggaatttagaganaaatalactatagattta 5640

[illegible][illegible]

Oy	1688	gataaaatctaacccacggcggaaaaaagtttgcatttcogcagtgacaagg	1747
Dd	613	ww	554
Oy	1748	ataaacctaacttaatttttgaanaaygaaccttctaactagaattaaacc	1807
Dd	553	ww	494
Oy	1808	gtatgtttttgttgccaatgacagccctcaacctgtagcgaatttttctcaa	1867
Dd	493	ww	434
Oy	1868	atataaattaggaattcaatgctactacaatagaagaacagctgaattacattt	1927
Dd	433	ww	374
Oy	1928	aattaagaacaaaattttgaaaaygtbataatttctaacaattataaataga	1987
Dd	373	ww	314
Oy	1988	tgcctataatgtatttccctagtgtccttaaataatttttatattagtataaac	2047
Dd	313	ww	254
Oy	2048	atataaccaataatagttcgtcgaaatccaatcatcattcataatttttgaaccta	2107
Dd	253	ww	194
Oy	2108	caaatcttaattatgtgcataaacaatgacabagaagttccaaaaaaatttgttaa	2167
Dd	193	ww	134
Oy	2168	cagaacttccaaatttttttttctatggaacagaaataacagatgaaaaacatttt	2227
Dd	133	ww	74
Oy	2228	gtctggaaatggaatgataataacatgaagcaattttaaaaaattataagcctat	2287
Dd	73	ww	14
Oy	2288	a 2288	
Dd	13	w 13	
RESULT 3			
ID.	AAF58254/C		
ID.	AAF58254 standard; DNA; 936 BP.		
XX	AAF58254;		
AC			
XX	24-APR-2001 (first entry)		
DT			
XX	Oligonucleotide D1875.		
DE			
XX	Electron-transfer group; ETM; mismatch; genotyping;		
KW	gene expression; ss.		
RW			
XX	Synthetic.		
OS			
XX	MO200107665-A2.		
PN			
XX	01-FEB-2001.		
PD			
XX	26-JUL-2000; 2000MC-US20476.		
PF			
XX	26-JUL-1999; 990S-0145695.		
PR	17-MAR-2000; 2000US-0190259.		
XX	(CLIN-) CLINICAL MICRO SENSORS INC.		
PA			
XX			
P1	Umek RM;		

XX WP1: 2001-159728/16.  
 DR  
 XX  
 PT Nucleic acids containing electron-transfer group, useful as labels in  
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on  
 PT a single surface -  
 XX  
 XX  
 PS Example 6; Page 127; 159pp; English.  
 PS  
 XX  
 CC The present invention relates to a composition comprising two nucleic  
 CC acids each containing an electron-transfer group (ETW) having  
 CC different redox potentials. The invention is used for electronic  
 CC detection of nucleic acids, especially of substitutions (mismatches)  
 CC and single-nucleotide polymorphisms, e.g. for genotyping,  
 CC monitoring gene expression.  
 CC  
 XX  
 SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match	3.6%;	Score 251;	DB 22;	Length 936;
Best Local Similarity	1.0%;	Pred. No. 1.2e-30;		
Matches	8;	Conservative 589;	Mismatches 184;	Indels 0;
			Gaps	0;

OY	1508	gcggtacatctcgcaacttaatttgcttctttaagtcataccttaaccttttt	1567
Dd	793	GGGCGCMMNN	734
OY	1568	gttagtgtgaatttaaagaagcgaaactgttaacattaacattatagtactag	1627
Dd	733	NNNGNNNNN	674
OY	1628	tatgtgatattcccaatatcatcacttggatgtttaacctaactctgttcctcc	1687
Dd	673	NN	614
OY	1688	gataaatatlaatcgcagagtaaaaaagttgtcttatcttcgcgafcgataag	1747
Dd	613	NN	554
OY	1748	ataaacctaagctttaatttttggaaaagtacaacctttaactcatagattaacc	1807
Dd	553	NNNG	494
OY	1808	gtaagtgttggccataatgaagccctcaacactgtgtagtcaaattttctcgaa	1867
Dd	493	NN	434
OY	1868	ataataataggaaatcataagctactacaatagaagaacagctgataltaacttt	1927
Dd	433	NN	374
OY	1928	aatttaagaacaaatttttggaaaagtataatcttaaccaaatataataaatga	1987
Dd	373	NN	314
OY	1988	tgccataatglatctccatgtgtcttaaacatttttttataattagttacaataac	2047
Dd	313	NN	254
OY	2048	attatgaaccaataatagttgtgaattccaatcctcataalatlttggaaatcta	2107
Dd	253	NNNC	194
OY	2108	caaatattaaatttagtccataaacaaagcagtagaagttccaaaaaaatttgtaa	2167
Dd	193	NN	134
OY	2168	cagaactccaaatttttttttatbgaaagaataacagatagaacataatttt	2227
Dd	133	NN	74
OY	2228	gttgctggaatggaagtagtatacatgaagcaaatttaaaaaattataagcctat	2287

```
Db 73 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 14
Oy 2288 a 2288
Db 13 W 13

RESULT 4
AAF58257/C
ID AAF58257 standard; DNA; 936 BP.
XX
AC AAF58257;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1954.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000MO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 3.6%; Score 251; DB 22; Length 936;
Best Local Similarity 1.0%; Pred. No. 1,2e-30;
Matches 8; Conservative 589; Mismatches 184; Indels 0; Gaps 0;

Oy 1508 ggcgttaaccatcgaacttaaatgcttcttaatagcattcaattattatt 1567
Db 793 GGGGGCXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 734
Oy 1568 gttagtgaattataagaagcgaactgtaacattacaattatattagactag 1627
Db 733 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 674
Oy 1628 tatgtatattccaatacatcattgagatgttaaacattacattgcttcctacg 1687
Db 673 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 614
Oy 1688 gtataaatatcatcagaggtaaaaaagtttgcctatttcgcgattcagaaag 1747
Db 613 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 554
```

```
Oy 1748 ataaacctacttaatttttgaaaatgtaaccttactacatagattaac 1807
Db 553 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 494
Oy 1808 gtaigtlttltgccaatagacagccttacaactgtgtagtcaattttctgca 1867
Db 493 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 434
Oy 1868 atataaataggaattcactatcatcagagaagaacgctgattacattt 1927
Db 433 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 374
Oy 1928 aatttaagacaaattttgaaaatgttataatttcaacaatatataataga 1987
Db 373 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 314
Oy 1988 tgcataatgtattcctatgttcttaaatattttttatattagttataaac 2047
Db 313 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 254
Oy 2048 attatgaaccaataatagttgtagaattcaaatatcctcatatatttttgaacta 2107
Db 253 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 194
Oy 2108 caattatattatagtcacataacatgcataagaagttccaaaaaatttgttaa 2167
Db 193 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 134
Oy 2168 cagaacttccaatttttttttattatgaaacaaataacagatagaactattt 2227
Db 133 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 74
Oy 2228 gttgtggaatggaagttagatatacatatagaacaattttaaatatataagcctat 2287
Db 73 WWWWXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 14
Oy 2288 a 2288
Db 13 W 13

RESULT 5
AAF58259/C
ID AAF58259 standard; DNA; 936 BP.
XX
AC AAF58259;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000MO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
```

PT a single surface -  
XX  
PS Example 6, Page 128; 159pp; English.  
XX  
CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an electron-transfer group (ETW) having  
CC different redox potentials. The invention is used for electronic  
CC detection of nucleic acids, especially of substitutions (mismatches)  
CC and single-nucleotide polymorphisms, e.g. for genotyping,  
CC monitoring gene expression.  
XX  
SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;  
  
Query Match 3.6%; Score 251; DB 22; Length 936;  
Best Local Similarity 1.0%; Pred. No. 1.2e-30;  
Matches 8; Conservative 589; Mismatches 184; Indels 0; Gaps 0;  
  
OY 1508 gcggttaaccatcgcgaacttaattgcttcttaatgcatcttaacattatt 1567  
DB 793 GGGGGCGMWM 734  
OY 1568 gttagtggaaatataaagcggaactgttaacattacattatagatactag 1627  
DB 733 WMMWM 674  
OY 1628 tatgtattatccaatacactctgagatgttaactaactctgttctctcag 1687  
DB 673 WMMWM 614  
OY 1688 gtataaatatcatcgcaggtataaaaagtgttgccttaatttcgcgatagaag 1747  
DB 613 WMMWM 554  
OY 1748 ataaactaagtacttaatttttgaagaatgaaccttactacatagaataacc 1807  
DB 553 WMMWM 494  
OY 1808 gtatgttttggtcgaataagacgctcacaacgtgatagtcaattttctgcaa 1867  
DB 493 WMMWM 434  
OY 1868 atataaataggaatcaatgctactatacagaagaacgctgatactattt 1927  
DB 433 WMMWM 374  
OY 1928 aatttaagcaaaacttctgaaaatgtataatcttcaaatattataaataga 1987  
DB 373 WMMWM 314  
OY 1988 tgcctataatgatttctctatgctcttaaaatattttttatattagtataaatac 2047  
DB 313 WMMWM 254  
OY 2048 atatagaacataatagttgtgattcaataatactccattatattttgaacta 2107  
DB 253 WMMWM 194  
OY 2108 caaatataatatttagtaacatacagatagaagaatccaataaatttctgtaa 2167  
DB 193 WMMWM 134  
OY 2168 cagaactccaattttttttttttagaagaagaataacagataagaactattt 2227  
DB 133 WMMWM 74  
OY 2228 gtctggaatggaagtagtaatacatttaagaatttaaaataataagcctat 2287  
DB 73 WMMWM 14  
OY 2288 a 2288  
DB 13 W 13

RESULT 6  
AAF58262/c  
ID AAF58262 standard; DNA; 936 BP.  
XX  
AC AAF58262;  
XX  
DT 24-APR-2001 (first entry)  
XX  
DE Oligonucleotide D2007.  
XX  
DE  
XX  
KW Electron-transfer group; ETW; mismatch; genotyping;  
KW gene expression; ss.  
XX  
OS Synthetic.  
XX  
PN W0200107665-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 26-JUL-2000; 2000MO-US20476.  
XX  
PR 26-JUL-1999; 99US-0145695.  
PR 17-MAR-2000; 2000US-0190259.  
XX  
PA (CLIN-) CLINICAL MICRO SENSORS INC.  
XX  
PI Umek RM;  
XX  
DR WPI; 2001-159728/16.  
XX  
PT Nucleic acids containing electron-transfer group, useful as labels in  
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on  
PT a single surface -  
XX  
PS Example 6, Page 128; 159pp; English.  
XX  
CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an electron-transfer group (ETW) having  
CC different redox potentials. The invention is used for electronic  
CC detection of nucleic acids, especially of substitutions (mismatches)  
CC and single-nucleotide polymorphisms, e.g. for genotyping,  
CC monitoring gene expression.  
XX  
SQ Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;  
  
Query Match 3.6%; Score 251; DB 22; Length 936;  
Best Local Similarity 1.0%; Pred. No. 1.2e-30;  
Matches 8; Conservative 589; Mismatches 184; Indels 0; Gaps 0;  
  
OY 1508 gcggttaaccatcgcgaacttaattgcttcttaatgcatcttaacattatt 1567  
DB 793 GGGGGCGMWM 734  
OY 1568 gttagtggaaatataaagcggaactgttaacattacattatagatactag 1627  
DB 733 WMMWM 674  
OY 1628 tatgtattatccaatacactctgagatgttaactaactctgttctctcag 1687  
DB 673 WMMWM 614  
OY 1688 gtataaatatcatcgcaggtataaaaagtgttgccttaatttcgcgatagaag 1747  
DB 613 WMMWM 554  
OY 1748 ataaactaagtacttaatttttgaagaatgaaccttactacatagaataacc 1807  
DB 553 WMMWM 494  
OY 1808 gtatgttttggtcgaataagacgctcacaacgtgatagtcaattttctgcaa 1867



```
Db 493 WWWWWW.....: 434
1868 attaatgaatcaatgctactactaataagaagaacagctggttaattt 1927
Db 433 WWWWWW.....: 374
1928 aattaaagacaaattttgaaaatgtataatttctaacaattataatga 1987
Db 373 WWWWWW.....: 314
1988 tgcctaatgtattcctcactgcttaaatattttttatatagttataaac 2047
Db 313 WWWWWW.....: 254
2048 attatgaaccataatagtgtggaatccaataatcctaataattttgaaatc 2107
Db 253 WWWWWW.....: 194
2108 caaatataatattagtaataaataagtagaagttccaaaaaatttggtaa 2167
Db 193 WWWWWW.....: 134
2168 cagaactccaattttttttttatggaacaaataacagatagaataattt 2227
Db 133 WWWWWW.....: 74
2228 gtgtggaatggaatgataataatcattagcaattttaaataatataagccat 2287
Db 73 WWWWWW.....: 14
2288 a 2288
13 W 13

RESULT 7
AAFS8255/c
ID AAF58252 standard; DNA; 938 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1876.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI. 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
```

```
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 3.6%; Score 251; DB 22; Length 938;
Best Local Similarity 1.0%; Pred. No. 1,2e-30;
Matches 8; Conservative 589; Mismatches 184; Indels 0; Gaps 0;

Oy 1508 ggtgtgacacatgcgaacttaattgcttcttaattgacttcaactatttt 1567
Db 793 GGGGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 734
Oy 1568 gtaagtgaatttaataagagcgaactgtatacataattatataatagactag 1627
Db 723 WWWWWW.....: 674
Oy 1628 tatgtatattccaataacttggatgtttaacttaacttcttctctcag 1687
Db 673 WWWWWW.....: 614
Oy 1688 gtataataatcatcagaggtlaaaaaagtttgcctatttcgcgagtgatgaag 1747
Db 613 WWWWWW.....: 554
Oy 1748 ataaactaatgacttaatttttgaaaatgtlaacccttactacatagataatcc 1807
Db 553 WWWWWW.....: 494
Oy 1808 gtaatttttgcctcaataatgacgcctctcaactgtgataagccaattttctgcaa 1867
Db 493 WWWWWW.....: 434
Oy 1868 attaatgaatgaattcaatgctactactaataagaagaacagctggaatatttt 1927
Db 433 WWWWWW.....: 374
Oy 1928 aattaaagacaaattttgaaaatgtataatttctaacaattataatga 1987
Db 373 WWWWWW.....: 314
Oy 1988 tgcctaatgtattcctcactgcttaaatattttttatatagttataaac 2047
Db 313 WWWWWW.....: 254
Oy 2048 attatgaaccataatagtgtggaatccaataatcctaataattttgaaatc 2107
Db 253 WWWWWW.....: 194
Oy 2108 caaatataatattagtaataaataagtagaagttccaaaaaatttggtaa 2167
Db 193 WWWWWW.....: 134
Oy 2168 cagaactccaattttttttttatggaacaaataacagatagaataattt 2227
Db 133 WWWWWW.....: 74
Oy 2228 gtgtggaatggaatgataataatcattagcaattttaaataatataagccat 2287
Db 73 WWWWWW.....: 14
Oy 2288 a 2288
13 W 13

RESULT 8
AAFS8252
ID AAF58252 standard; DNA; 936 BP.
```









QY	2084	tcacatataatttttggaaatccacaacattataatttagtcataataacatgataga	2143
Db	541	.....	600
QY	2144	aagttccaaaaaaatttgytatacagaactccaattttttttttagagacaag	2203
Db	601	.....	660
QY	2204	aaataacagatgaaaactatttgttcfgaaatggaagtagtaataacattaaagcaa	2263
Db	661	.....	720
QY	2264	ttttaaataattatagaacctaacgcgtcaagtatgtatctagtagtgaarta	2323
Db	721	.....	780
QY	2324	ataat	2328
Db	781	.....	785

RESULT 13  
AAFS8255 standard: DNA: 938 BP.  
AAFS8255;  
AAFS8255;  
24-APR-2001 (first entry)  
Oligonucleotide D1876.

XX	Synthetic.
OS	
XX	WO200107665-A2.
PN	
XX	01-FEB-2001.
PD	
XX	26-JUL-2000; 2000MO-US20476.
PF	
XX	26-JUL-1999; 99US-0145695.
PR	
XX	17-MAR-2000; 2000US-0190259.
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
PI	
XX	Umek RM;
DR	
XX	WPI; 2001-159728/16.
XX	
PT	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX	a single surface
XX	
PS	Example 6; Page 127; 159pp; English.
XX	
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETM) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,
CC	monitoring gene expression.
XX	
XX	
SO	Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
	Query Match
	Best Local Similarity 3.6%; Score 250.6; DB 22; Length 938;
	Matches 4; Conservative 596; Mismatches 185; Indels 0; Gaps

Db 1 www..... 60  
Qy 1604 tacaattatataatagatactagatgtatattccaataataacttggatgtta 1663  
Db 61 ..... 120  
Qy 1664 aacttaacttcttctcctacgataaaatlaatcaacgaggtaaaaagtttg 1723  
Db 121 ..... 180  
Qy 1724 tcttatttcgagatgaagataaacctaagacttaattttttgaaaatgttaac 1783  
Db 181 ..... 240  
Qy 1784 ccttactacatagataatgaatgctgtttgtgtccataatgaagcctctaac 1843  
Db 241 ..... 300  
Qy 1844 tctgataagtcataatttctcgcataataatlaagaaatcaatgctactcaataga 1903  
Db 301 ..... 360  
Qy 1904 agaagacgctgagatgaatlaaagacaanaattttgaaaaatgtataatt 1963  
Db 361 ..... 420  
Qy 1964 tctacaataataataatagatgctcctaatagtatttctcctaataatt 2023  
Db 421 ..... 480  
Qy 2024 tttttatatttgataataatcatatgaaccaataaagttgtgaattcaaatc 2083  
Db 481 ..... 540  
Qy 2084 tccattaaatttttgaatactcaaaattataattagtcataataacatgcataga 2143  
Db 541 ..... 600  
Qy 2144 aagttccaaaaaatttgttaacagaaactccaatttttttttataagaaag 2203  
Db 601 ..... 660  
Qy 2204 aatlaacagatagaataactatttgtgtgaaatgagagatgaataataacaa 2263  
Db 661 ..... 720  
Qy 2264 tttaaaaaatataaagcctatacgcgtcaaaatgtatctagtagtgaatta 2323  
Db 721 ..... 780  
Qy 2324 ataact 2328  
Db 781 ..... 785

RESULT 14  
ID AAT85307 standard; DNA: 4937 BP.  
AC AAT85307;  
XX  
DT 14-APR-1998 (first entry)  
XX  
DE Arabidopsis thaliana cytochrome P450-type hydroxylase DNA.  
XX  
KW Cytochrome P450-type hydroxylase; identification; brassinosteroid;  
KM brassinosteroid inhibitor; modified plant; recombinant production;  
XX  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
FH Key Location/Qualifiers  
FT CDS 968..4773

FT FT /\*tag= a  
FT FT /product= cytochrome\_P450-type\_hydroxylase  
FT FT 1484..1679  
FT FT /\*tag= b  
FT FT /number= 1  
FT FT 1830..1916  
FT FT /\*tag= c  
FT FT /number= 2  
FT FT 2166..3902  
FT FT /\*tag= d  
FT FT /number= 3  
FT FT 3990..4083  
FT FT /\*tag= e  
FT FT /number= 4  
FT FT 4163..4247  
FT FT /\*tag= f  
FT FT /number= 5  
FT FT 4355..4445  
FT FT /\*tag= g  
FT FT /number= 6  
FT FT 4577..4673  
FT FT /\*tag= h  
FT FT /number= 7  
XX XX  
PN MO9735986-A1.  
XX  
PD 02-OCT-1997.  
XX  
XX 27-MAR-1997; 97MO-EP01586.  
PF  
XX 27-MAR-1996; 96US-0622166.  
PR  
XX (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PA  
XX Altmann T, Koncz C, Mathur J, Szekeres MA;  
XX  
XX WPI: 1997-489649/45.  
DR P-PSDB; AAM27153.  
XX  
XX  
XX New isolated plant cytochrome P450-type hydroxylase gene - used to  
PT identify substances acting as brassino-steroid(s) or brassinosteroid  
PT inhibitors for the production of modified plants  
XX  
XX Claim 1; Pages 48-53; 77pp; English.  
PS  
XX  
XX The present sequence encodes Arabidopsis thaliana cytochrome  
CC P450-type hydroxylase. The hydroxylase can be used to identify  
CC brassinosteroids or brassinosteroid inhibitors, useful to produce  
CC plants with modified physiological and/or phenotypic  
CC characteristics. The modified plants may show, e.g. stimulation of  
CC growth, increased cell elongation, increased wood production,  
CC accelerated seed germination at low temperatures, an increase in  
CC dry weight, repressed anthocyanin production during growth in light  
CC and/or inhibited de-etiolation which is induced, e.g. by cytokinin,  
CC in the dark or an increase in stress tolerance. The hydroxylase or  
CC its coding sequence can also be used for the recombinant production  
CC of compounds, e.g. teasterone.  
XX  
XX Sequence 4937 BP; 1526 A; 864 C; 870 G; 1677 T; 0 other;  
SQ

Query Match 1.3%; Score 92.8; DB 18; Length 4937;  
Best Local Similarity 53.1%; Pred. No. 4.1e-06;  
Matches 225; Conservative 0; Mismatches 192; Indels 7; Gaps 1;

Qy 3501 caggtatgtaagaataatagatcgactgtttggaacacgaagatgcatcagctga 3560  
Db 1162 ccggtacggtcgtttcattacatgaacgaaggaagcctcttgaaatgtagttatcc 1221  
Qy 3561 tcttgactaataagatacatatacaaaacgaaggaagcctcttgaaatgtagttatcc 3620  
Db 1222 ccggaacgaacgcggtttgttcttcaagaacgaagggaagccttttgagtgcttctatcc 1281

OY	3621	tagaagatagtgctgggatttccttgcggaacggtccgatgcgtttcgtttgttgcacgtca	3680
Db	1282	tgtctcatctttnaaccttcttggcggaacacctctgcgtctctaagaagtctttgca	1341
OY	3681	tagagataatagaagaatgatctgcgttaactctttaagtcacgcagctcttagaactatct	3740
Db	1342	taaaagtatgcactctctccaccatgagctcttgcataattccaatacataaaggacactt	1401
OY	3741	acttaagaatgcttgcagagacaataccttgttcttgcattcttcgtgcacaaaactctat	3800
Db	1402	cattgcttgatatattgaccggttagtcggtttaactctgtactcttg-----tctctc	1454
OY	3801	tttctctgcctccaagacgagccaaaaagttttattarcttcttatttgcctaatt	3860
Db	1455	gtgtctcccaccatggaagaagccaaaaggtaaccaaaaaaatcttgcattacaanaac	1514
OY	3861	tttttgccttaagaactctltagagtlctctaacttttttttttaattgaacagtttaagt	3920
Db	1515	attatattattatttaatttcgcgtcttcactatagtttttttataataaataataaat	1574
OY	3921	tttaa	3924
Db	1575	aaaa	1578

XX	AAAF58238	standard; DNA; 244 BP.
XX	AAAF58238;	
XX	24-APR-2001	(first entry)
XX	Oligonucleotide	D1250:D1102.
XX	Electron-transfer group; ETM;	mismatch; genotyping;
XX	gene expression; ss.	
XX	Synthetic.	
XX	WO200107665-A2.	
XX	01-FEB-2001.	
XX	26-JUL-2000;	2000WO-US20476.
XX	26-JUL-1999;	99US-0145695.
XX	17-MAR-2000;	2000US-0190259.
XX	(CLIN-) CLINICAL MICRO SENSORS INC.	
XX	Umek RM;	
XX	WPI: 2001-159728/16.	
XX	Nucleic acids containing electron-transfer group, useful as labels in	
XX	hybridization assays, e.g. for genotyping, allowing repeat analyses on	
XX	a single surface	
XX	Example 4; Page 120; 159pp; English.	
XX	The present invention relates to a composition comprising two nucleic	
XX	acids each containing an electron-transfer group (ETM) having	
XX	different redox potentials. The invention is used for electronic	
XX	detection of nucleic acids, especially of substitutions (mismatches)	
XX	and single-nucleotide polymorphisms, e.g. for genotyping,	
XX	monitoring gene expression.	
XX	Sequence 244 BP; 19 A; 9 C; 12 G; 10 T; 194 other;	

Query Match	1.38;	Score 88.2;	DB 22;	Length 244;
Best Local Similarity	8.9%;	Pred. No. 1.7e-05;		

[illegible]

```
Search completed: September 29, 2001, 11:29:26
Job time: 14750 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2001, 06:49:51 : Search time 8875.91 seconds  
(without alignments)  
12003.485 Million cell updates/sec

Title: US-09-502-426-1  
Perfect score: 6888  
Sequence: 1 atgtgggtattatattgttg.....cctagtcagcgccaaagctt 6888

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_ba3:\*  
4: gb\_in1:\*  
5: gb\_in2:\*  
6: gb\_in3:\*  
7: gb\_on:\*  
8: gb\_ov:\*  
9: gb\_pat1:\*  
10: gb\_pat2:\*  
11: gb\_ph:\*  
12: gb\_pl1:\*  
13: gb\_pl2:\*  
14: gb\_pl3:\*  
15: gb\_pl4:\*  
16: em\_ba1:\*  
17: em\_ba2:\*  
18: em\_fun:\*  
19: em\_htgo\_hum:\*  
20: em\_htgo\_inv:\*  
21: em\_htgo\_rod:\*  
22: em\_htg\_hum1:\*  
23: em\_htg\_hum2:\*  
24: em\_htg\_hum3:\*  
25: em\_htg\_hum4:\*  
26: em\_htg\_hum5:\*  
27: em\_htg\_hum6:\*  
28: em\_htg\_hum7:\*  
29: em\_htg\_hum8:\*  
30: em\_htg\_inv1:\*  
31: em\_htg\_inv2:\*  
32: em\_htg\_other:\*  
33: em\_htg\_rod:\*  
34: em\_hum1:\*  
35: em\_hum2:\*  
36: em\_hum3:\*  
37: em\_hum4:\*  
38: em\_hum5:\*  
39: em\_hum6:\*  
40: em\_hum7:\*  
41: em\_in:\*  
42: em\_on:\*  
43: em\_or:\*

44: em\_ov:\*  
45: em\_pat:\*  
46: em\_ph:\*  
47: em\_pl:\*  
48: em\_ro:\*  
49: em\_sts:\*  
50: em\_sy:\*  
51: em\_un:\*  
52: em\_v1:\*  
53: gb\_sts1:\*  
54: gb\_sts2:\*  
55: gb\_sts3:\*  
56: gb\_sy:\*  
57: gb\_un:\*  
58: gb\_v1:\*  
59: gb\_v12:\*  
60: gb\_htg1:\*  
61: gb\_htg2:\*  
62: gb\_htg3:\*  
63: gb\_htg4:\*  
64: gb\_htg5:\*  
65: gb\_htg6:\*  
66: gb\_htg7:\*  
67: gb\_htg8:\*  
68: gb\_htg9:\*  
69: gb\_htg10:\*  
70: gb\_htg11:\*  
71: gb\_htg12:\*  
72: gb\_htg13:\*  
73: gb\_htg14:\*  
74: gb\_htg15:\*  
75: gb\_htg16:\*  
76: gb\_htg17:\*  
77: gb\_htg18:\*  
78: gb\_htg19:\*  
79: gb\_htg20:\*  
80: gb\_htg21:\*  
81: gb\_htg22:\*  
82: gb\_htg23:\*  
83: gb\_htg24:\*  
84: gb\_htg25:\*  
85: gb\_pr1:\*  
86: gb\_pr2:\*  
87: gb\_pr3:\*  
88: gb\_pr4:\*  
89: gb\_pr5:\*  
90: gb\_pr6:\*  
91: gb\_pr7:\*  
92: gb\_pr8:\*  
93: gb\_pr9:\*  
94: gb\_rot1:\*  
95: gb\_rot2:\*  
96: gb\_in4:\*  
97: gb\_pr10:\*  
98: em\_ba3:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
C 1	6753.6	98.0	84196	14 ATT3A5	AL132979 Arabidops
2	4818	69.9	4818	12 AF044216	AF044216 Arabidops
3	134	1.9	100269	13 AF18022	AF18022 Arabidops
4	133.2	1.9	104992	60 AC005504	AC005504 Plasmodiu
5	133.2	1.9	169546	60 AC004157	AC004157 Plasmodiu
C 6	130.8	1.9	104992	60 AC005504	AC005504 Plasmodiu
7	130.8	1.9	169546	60 AC004157	AC004157 Plasmodiu
8	129.4	1.9	14867	4 AE001398	AE001398 Plasmodiu

9	128	1	9	106142	12	AC025417	AC025417 Genomic s
10	127.8	1	9	86827	96	PFMAL3P5	AL034556 Plasmodiu
11	125.4	1	8	86827	96	PFMAL3P5	AL034556 Plasmodiu
12	124.4	1	8	4601	6	DMU11584	U11584 Drosophila
13	124.4	1	8	19517	6	DMU37541	U37541 Drosophila
14	123.4	1	8	67970	96	PFMAL1P3	AL031746 Plasmodiu
15	122.6	1	8	4601	6	DMU11584	U11584 Drosophila
16	122.6	1	8	19517	6	DMU37541	U37541 Drosophila
17	118	1	7	67970	96	PFMAL1P3	AL031746 Plasmodiu
18	118	1	7	84196	14	ATT3A5	AL132979 Arabidops
19	113	1	6	14867	4	AE001398	AE001398 Plasmodiu
20	107.8	1	6	152209	92	HS1108D11	AL034419 Human DNA
21	105.6	1	5	176174	86	AC007483	AC007483 Homo sapi
22	103	1	5	130117	85	AC004907	AC004907 Homo sapi
23	103	1	5	149733	77	AC087568	AC087568 Pan trogl
24	102.6	1	5	15643	5	AF149768	AF149768 Bombyx mo
25	101.4	1	5	178087	85	AC005089	AC005089 Homo sapi
26	101.2	1	5	15421	96	PRCOMPTRA	X95275 P.falciparu
27	100.8	1	5	152209	92	HS1108D11	AL034419 Human DNA
28	100.8	1	5	188450	79	AL158209	AL158209 Homo sapi
29	99.6	1	4	180741	75	AC073846	AC073846 Homo sapi
30	99.4	1	4	1867	14	MSCA1J23	AJ223323 Saccharom
31	99.4	1	4	175223	85	AC005410	AC004617 Homo sapi
32	99.4	1	4	310779	60	AC005140	AC005140 Plasmodiu
33	99.9	1	4	185603	74	AC073136	AC073136 Homo sapi
34	98.4	1	4	2290	6	MIDYTRN	X05915 D. yakuba m
35	98.4	1	4	16019	6	MIDYTRN	X03240 Drosophila
36	98.2	1	4	12029	4	AE001427	AE001427 Plasmodiu
37	97.4	1	4	62268	92	HS1178121	AL109852 Human DNA
38	97.2	1	4	159781	69	AC025018	AC025018 Homo sapi
39	97	1	4	159593	85	AC004832	AC004832 Homo sapi
40	96.8	1	4	195969	75	AC074354	AC074354 Oryza sat
41	96.6	1	4	234112	96	PFMAL4P2	AL035475 Plasmodiu
42	96.4	1	4	242893	83	CEY53C12	Z92859 Caenorhabet
43	96	1	4	146285	85	AC005083	AC005083 Homo sapi
44	95.8	1	4	155012	78	AL133476	AL133476 Homo sapi
45	95.6	1	4	156060	60	AC004153	AC004153 Plasmodiu

## ALIGNMENTS

RESULT 1	ATT3A5	84196 bp	DNA	26-JAN-2000
LOCUS	Arabidopsis thaliana DNA chromosome 3, BAC clone T3A5.			
DEFINITION	AL132979			
ACCESSION	AL132979.2			
VERSION	GI:6782244			
KEYWORDS	thale cress.			
SOURCE	Arabidopsis thaliana			
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 84196)			
AUTHORS	Bloeker, H., Mewes, H.W., Lemcke, K., Mayer, K.F.X., Quetier, F. and Salanoubat, M.			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 84196)			
AUTHORS	EU Arabidopsis sequencing, project.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-JAN-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Marcel Salanoubat and Francis Quetier, Groupement d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue Gaston Cremieux, BP191, 91006 Evry Cedex, France; http://www.genoscope.cns.fr			
COMMENT	On Jan 27, 2000 this sequence version replaced gi:6434247. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/. Location/Qualifiers			
FEATURES				

source	1. .84196	/organism="Arabidopsis thaliana"
		/variety="Columbia"
		/db_xref="taxon:3702"
		/chromosome="3"
LTR	2983. .3338	
		/note="355 bp LTR"
misc_feature	3339. .7564	
LTR	7565. .7922	
		/note="LTR-transposon"
gene	12339. .13197	
		/gene="T3A5.10"
		/join(12339. .12665,12742. .12803,12888. .13055,13125. .13197)
CDS	/gene="T3A5.10"	
		/note="similarity to cyclin-dependent kinase inhibitor (ICK1), Arabidopsis thaliana, EMBL:AF094772.1"
		/codon_start=1
		/product="putative protein"
		/protein_id="CAB62432.1"
		/db_xref="GI:6561966"
		/translation="MAVRRREEDVVEENGVTITVYKRRKEEVDLVESRIILPCV OATNRGIVARNAGASSTSVYVRRRSDPPVEOCOTREEDSSVSCSTSEKSKRR IEPVDEENKGDREYETSWITDILKSESNMDSVAVEDVESRRRLRSKHETV KEAELEDFQVAKEDLRNKLLECSMTNFDPEKDEPLGGRTIEWKLP"
exon	12339. .12665	
		/gene="T3A5.10"
		/number=1
Intron	12666. .12741	
		/gene="T3A5.10"
		/number=1
exon	12742. .12803	
		/gene="T3A5.10"
		/number=2
Intron	12804. .12887	
		/gene="T3A5.10"
		/number=2
exon	12888. .13055	
		/gene="T3A5.10"
		/number=3
		/number=3
		/gene="T3A5.10"
		/number=3
		/gene="T3A5.10"
		/number=4
exon	13125. .13197	
		/gene="T3A5.10"
		/number=4
exon	15858. .16470	
		/gene="T3A5.20"
		/complement(join(15858. .15901,16014. .16470))
		/note="similarity to predicted proteins, Arabidopsis thaliana"
		/codon_start=1
		/product="putative protein"
		/protein_id="CAB62433.1"
		/db_xref="GI:6561967"
		/translation="MARLEHNSMNGARISFSENFVEIRSEKSNKNNINSRSSFS MSPADAFSVTDMSIAPADEIFLKGKILPFKETSVRTLGEEILTEEGSMWDGTF SRRLPILSSSFSTKGTWRELLGLKRTVRSKTKDVNREVLISQDKIISGNVATREC QVADPR"
		/complement(15902. .16013)
		/number=1
		/complement(16014. .16470)
		/number=2
gene	18303. .19931	
		/gene="T3A5.30"
		/complement(18303. .19931)
exon	/number=1	
		/complement(18303. .19931)
CDS	/codon_start=1	
		/product="scarce row-like 7 (SCL7)"
		/protein_id="CAB62434.1"



/db\_xref="GI:6561968"  
/translation="MAYMTDSCNIMAIQAOLIKKQOQOQSOHQOQEEQOQEPMPN  
PSFGTLPSSGSPFQVNDPGFHPHLEHMOHAAVASEEPDSEWESIINGDS  
QTPDPPIVGHDPVSPSPRLSAPSYLNRYNKDQSAOOLPPASPAIATSPSPSQ  
HPPPPPPPPDPLNPIFKAIIDHAKRPETKPTLLIKESVSESGDPIORVGTFAE  
ALSHKEEPSSSSSSSSSEDFILSYKTLNACPSKPAHLTANQALILEANQNNHII  
VDIGIFQGIOMSLQALATRSSGKPRISGIPAPSLGSDPGSLIAGNRLRDA  
AIDLNFEFVPLTPIQLNGSSFRVDEVLVYFMLELYKLDEPATTYGTALRA  
RSLNPRIVTGEYVSLNRYEPANRYKNSLRFSYSAVPESLEPNDRDKRLRYEVL  
EGRRIMDLVSDDDNKPCTRFGLMEKEQRMVLEKAGFEPAKPSNVAVSQAKLLIM  
NYSITLVLSESPGFIISLANNVPLLTIVSMR"  
26093..28999  
/gene="T3A5.40"  
/complement(join(26093..26338,27091..27200,27308..27386,  
27478..27570,27723..28037,28136..28288,28374..28688,  
28779..28999))  
/note="Contains Cytochrome P450 cysteine heme-liron ligand  
signature AA455-464;Prokaryotic membrane lipoprotein lipid  
attachment site AA320-330"  
/codon\_start=1  
/product="steroid 22-alpha-hydroxylase (DWF4)"  
/protein\_id="CAB62435.1"  
/db\_xref="GI:6561969"  
/translation="MEETEHHTLPLLLPSLISLILLLILKRNKRRTFENLPGRKS  
GMPFLGETIGYIKPYATTLGDEMOCHSKYGYKLYRSNLGEPITVSADAGLNFILO  
NEGRLFECSYPSRISGIDLGKMSMLVIGMDHMRDMSISLNPISARLRTLLKQVEH  
TLFLYDSMOQNSIFGSADDEAKKTFENIMAKHMSMDPEEETBOLKEVYTFEMGYVS  
APLMPGTATYHKAQSRATLTKFERMEERKLDIKEDDEEVEVKDEDEENKSDH  
VKQRDTDDLLGWLKHNSLSTEQIIDLILSLFAGHETSVAIALAIFPLQAPKAV  
EELREHEILARAKKESESELMDYDKKDFQCVINETLRGNVRFYLRKALKY  
RYKGYDIPSGMKVLPVISAHLNDRDNLFPNMQOONGASSSGSGFTWGN  
NYMFGGGRCLAGSELALEMAVEIHLVLKFWELAEDDKPEAFVDFPNPLPR  
VSRLL"  
exon  
/complement(26093..26338)  
/number=1  
/complement(26339..27090)  
exon  
/number=1  
/complement(27091..27200)  
intron  
/number=2  
/complement(27201..27307)  
exon  
/number=2  
/complement(27308..27386)  
intron  
/number=3  
/complement(27387..27477)  
exon  
/number=3  
/complement(27478..27570)  
intron  
/number=4  
/complement(27571..27722)  
exon  
/number=4  
/complement(27723..28037)  
intron  
/number=5  
/complement(28038..28135)  
exon  
/number=5  
/complement(28136..28288)  
intron  
/number=6  
/complement(28289..28373)  
exon  
/number=6  
/complement(28374..28698)  
intron  
/number=7  
/complement(28699..28778)  
exon  
/number=7  
/complement(28779..28999)  
gene  
38307..41323  
/gene="T3A5.50"  
/complement(38307..38796)  
exon  
/number=1  
/complement(join(38307..38796,38877..39090,39189..39270,  
40180..40242,40337..40401,40489..40564,40655..40734,  
40805..40919,41325..41323))  
/note="Contains Eukaryotic putative RNA-binding region  
RNP-1 signature AAI/9-186"  
/codon\_start=1

/product="U1 snRNP 70K protein"  
/protein\_id="CAB62436.1"  
/db\_xref="GI:6561970"  
/translation="MGDSGPPFLRNPAAYQARAKYONRANVQLKMGOSHPTGLTN  
NLKLFEPREPLEKPPPERKPCPPYGMQFVSNFEPDDEPAPKPEVLEPSQR  
ERLHKLEGEVEKAAEDLKRPNDPNMATGDPYKTLFYSRLYESSEKIRKEFS  
YGPTRVHLVTDQITNPKGVALEYMHTDMKAAAYQAOQOKIDGRVIVYERGT  
VPMNRPRLRGIGSTSRVGGEEVSGOQGRGTSOESRSPREEKSRKEKERE  
RSRLSHEQREERSRDRPRDDKHRRDQGRDPRDRSDRDRDRDRDRDRDR  
GRDRTSHDRDRDRSRKREKRDYEGGEYHEGSGRSREDAEYRGPEETRGTYEDQGD  
TDRTSHRYDMEEDFRERYERKSRSESEYVR"  
26093..28999  
/gene="T3A5.40"  
/complement(38877..39090)  
exon  
/number=2  
/complement(39091..39188)  
intron  
/number=2  
/complement(39189..39270)  
exon  
/number=3  
/complement(39271..40179)  
intron  
/number=3  
/complement(40180..40242)  
exon  
/number=4  
/complement(40243..40336)  
intron  
/number=4  
Query Match 98.0%; Score 6753.6; DB 14; Length 84196;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 6864; Conservative 0; Mismatches 19; Indels 9; Gaps 9;  
OY 1 atgtggtatataatgttggttggttggttggtgacacataataatcgtgttcgtg  
|||||  
Db 32139 ATGTGGGATATATATGTGGGTTGGGTTGAGTCACATATTAATTCGTTCTGTGT  
OY 61 tatctgtcacatgattgagttggttccaatlttgatccaagataataatc  
|||||  
Db 32139 TATCTGTTCACATGATTGATTGTTGTTCAATTTGATCCAAAGATAATTAATAT  
OY 121 aaatcatttaaatattcaagtaataataatccttaccattgattgtttataaaa  
|||||  
Db 32079 AAAATTCATTTAAATATTTCAAGTAATTAATTAATTCATTTGATTAACAA  
OY 181 aatcctactcttggtatagaaataatgagttggaatttaataataaagaaa  
|||||  
Db 32019 AATATCTATCTTTGATATATGAAAAATATGAGTTTGCAATTTAATTAAGGAAA  
OY 241 taatcgattccattggttgattacacagttgaatttctgttcttctgttataatg  
|||||  
Db 31959 TAATCGATTCCATTTGTTGATTAACAGTTAAAGTTTGTGTTGTTATATGT  
OY 301 atatgagtaaatcaaaagagatattgtaagtgtaaacatatttctgttataaccca  
|||||  
Db 31899 ATATGATTAATCAAAAAGACATTAATTAATTAATTAATTAATTAATTAATTA  
OY 361 aaaaaaaac  
|||||  
Db 31839 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA  
OY 421 aggttatttgatcaataatcatgcatcatttcttgattactataagaatttcttac  
|||||  
Db 31780 AGGTTTATTTGATCATATTAATTCATGATCATTTCTTTATTAATTAATTAATTA  
OY 481 caatcaaaatttgcattcatatctctgattataataataatcggttggaatcc  
|||||  
Db 31720 CAATTAATAATTCGAATTCATATCTCTGATTATTAATTAATTAATTAATTAATTC  
OY 541 gttatcgatcacatccaatcatgattatctgtgtctaatccagaataattataaca  
|||||  
Db 31660 GTTATTCATCATCTCATCATGATTATGATCTTGCTGCTAATCCAGCAAAATTAACA  
OY 601 agagattatgagaaaacccaagaataaagaagaagagtagtgaccatggaatg  
|||||  
Db 31600 AGAGATTGAGAAAAAAGCAAAATAAAGGAAAGAGTAGACCCATGAGATG 31541

QY	661	tgataataatccaagagaaataagagatgaacaaccaaaagttgtggaataatggtccct	720
QY	31540	tgaaatatttatcaaaagagaataagagatgacacacaaaagttgtggaataatggtccct	31481
QY	721	ggcaagttctctccaacataatcacaccatttggattcttcgtgaattctgttaaa	780
Db	31480	ggcagcttctctcaacaaatcattgcaccatttggattcttcgtgaattctgttaaa	31421
QY	781	tttgagataagatctggaaaaaatcttatttgttagctgaatcgaatatagttcca	840
Db	31420	ttttggcattaacattgtgaaaaattttttttttgttagctgattcaaatattatgttcca	31361
QY	841	ggattttgataatctctgctttaagaaatcttgcctctctcttctgttcgttccct	900
Db	31360	ggattttgcataatctctgctttaaagcatttttgccttcttctttttgttgcgtttcct	31301
QY	901	taactatataatcgcggagataatgatatacaagatgatatacacaaacattgtctgg	960
Db	31300	taactatataatattcgcggaattattgatattacaaatgatattatcaaaaacaaattgtctgg	31241
QY	961	accattttgaataaacttttctcaaaccttagcgggaaccttggacttgaaccttaata	1020
Db	31240	accattttgaataaacttttctcaaaccttagcgggaaccttggacttgaaccttaata	31182
QY	1021	cgattttaaagcgtcaactgcttggagattactagataaagcaaaagacccgttcaagc	1080
Db	31181	cgatttttaaagcgtcactgcttggagattactagataaagcattaaagcattaaagccgttcaagc	31122
QY	1081	tatttatacaagttacaacatcgaaatagcttgaatccctttagaaaaatttggaaata	1140
Db	31121	ttttttatatacmaagttacaaacgaattatnagcttgaattccctttacaaaaattttggaaatta	31062
QY	1141	ccggtgttatgttaataatagattagtggtgaacaaatgatttaacaaatagtgta	1200
Db	31061	ccggtgttatgttaataatnagatttagtgctgataacaaatgtttatcatattagtgctca	31002
QY	1201	acatabaataatcccttaacagaaaaaacaacttaagagaagttaacatatcatatat	1260
Db	31001	acattatnactaatttccctttacaaaaaaacaaacttaagagaagttaacatatcatatat	30942
QY	1261	gggtatgcataaccttctacgtatctatactaaagacaaagaatagtattagtatgc	1320
Db	30941	gggtatgcataaccttcttctacgtatctatnactaaagacaaagaatagtattagtatgctc	30882
QY	1321	gataaatagaattcacacacgcgtgtgaataatataatggagaccgatgttcaacatctgca	1380
Db	30881	gatttaatagaatttccacacgcggtgtgaattatgtggagaccgatgtttacgatcactgcga	30822
QY	1381	atacatctcttggtgtgtcaacaataaaaaaacaagaagaagaagaagaatcttct	1440
Db	30821	aatatcttcttggtgtgtgtctacacaaatataaaaaaacaagaagaagaagaagaatcttct	30762
QY	1441	cttggatctcatcaatgatacttaaaatgataagatcttcttgggttaagatttgcgaagtc	1500
Db	30761	cttggatttccattcattgatttcttaaaatgcatnagattcttgggtttacagatttgcgaagtc	30702
QY	1501	cttacaagcgtgttaacacatctgcaactataatgtgcttctcttaatgtcatcttaaat	1560
Db	30701	cttacaagcgtgttaacacatctgcaactataatgtgcttctcttaatgtcatcttaaat	30642
QY	1561	atttattgttaattggaaattataagaagcgaacttgttaacattacatatattatattag	1620
Db	30641	atttattgttaattggaaatttttaataagacggaacttgttaacatttttaattattag	30582
QY	1621	atactagatctggaattatccaatacaatacttggatcttgaatcttaaacattctgtct	1680
Db	30581	atactagatctggaattatttccaaatnactacatttggattgttttaaacatttattctgtct	30522
QY	1681	tcctacggtataataattaaatcaatcgaggttaaaaaagtttgccttatcttcgcgatgc	1740
Db	30521	tcctatc-gattataaatttatcatctgaggttaaaaaagttttgtctttttttgcgagatgc	30463

QY	1741	atgaaggaataaacttaagcttaatttttttggaaaagttaacccctttactccatagtc	18000
Db	30462	ATGAAGGATTAACCTAATGACTTAATTTTTTGGAAATATGTAACCTTTACTCATAGATT	30400
QY	1801	aattaccgatgtcttctgttgcataatgacagccctctcaactgtgtagtcaatttc	18600
Db	30402	AATTACCGTATGTTTTGTTGGCATATGACACCGCTCTACACGTGTGATGCAATTTTT	30340
QY	1861	tcgtcaaatatlaaataggaaatccaatgctactcaatagaagaagaacagctgagatc	19200
Db	30342	TCTGCAAAATATTAATTAAGAAATTCATGCTACTATCAATTAAGAAACAGCTGAGATT	30280
QY	1921	acatttcaattlaaagacaaactttctgaaaaatgtctaacttctcaaatatatttaa	19800
Db	30282	ACATTTAATTTAAAGCAAAATTTTTGAAAAATGTTAATTTCTACAAATATATTAA	30222
QY	1961	aataatgacctataatgatatcttcctaagtccttaaatatttttctataattagta	20400
Db	30222	AATATGATGCTATAAAGATTTCTCATGTGCTCTTAAAAATTTTTTTAATATTAGTTA	30160
QY	2041	taaatactatagacccaataatagtttggtagtcaaatatctccatcaattatttctg	21000
Db	30162	TAAATACATTGTGACCAATATATAGTTGGTAATTAAATCTCCATTAATTTTTTTC	30100
QY	2101	aaactcaaaattlaataatttagtccaataacatgcaatagaagcttccaaaaaaat	21600
Db	30102	AAATCTACAAATTTATTAATTTTATAGTCAATTAACATGCTATGAAGTTCCAAAAATTT	30040
QY	2161	tgttgaacagaacttccaattttttttttttttatgaaacagaataatcagatagaaa	22200
Db	30042	TTGTTAATACGAATCTTCCAAATTTTTTTTTTTATGGAACAAAGAAATATACATAGAAA	29980
QY	2221	ctatttctgttggaaatggaagtagtaataacatagaacaaatttlaaaaatata	22800
Db	29982	CTATTGTGTGGAAGTAGAAGTAGTAATATATACATTAAACCAATTTTAAAAATATATA	29922
QY	2281	agcctatacgcgtccaagaatgatatctagtagtgtaataataatgcatgtgcat	23400
Db	29922	AGCCTATACGCGCTCAAGATATGTATTCTAGTAGTGTAATTAATTAATGACATGTCGAT	29860
QY	2341	tcgaatitggacacaaatgaaaaacggaattlaaatataacttlaaataataaataat	24000
Db	29862	TCGAGATTGGGACACACATGAATAACGGAATTTAAATATTAACTTTAAATATAATAAAT	29800
QY	2401	ttgagtaaatgcttctcgaactatcgaggggcaaaaaagacaatgccaaagctac	24600
Db	29802	TTTGAGTAAATGTGTTTTCTCATCTATTGAGGGGCAAAAAACACATGCAAAATCTCTAC	29740
QY	2461	gggtttgactgtccagttctgtaataatcctaatactgctcttggacgcagcgtctg	25200
Db	29742	GGGTTTGACTGTCCAGTTCGGTAATATCTAATTAACCTGTCTTTGACCGCACGCTCTGTG	29680
QY	2521	taggggtccctctgaacttctcaactgtctcaccctacacgtagcccaacctttccca	25800
Db	29682	TAGGGGTCTCTTGACATTTTTCACTGTCTACCCCTACGAGACCCACCCCTTTTCCCA	29620
QY	2581	tatccataaggatlaattcttgaaaatcccaattlaaaccgatctgagacgtagccgactcc	26400
Db	29622	TATCTCTAAGGGTAAATTTTGGAAATTCCTCAATTTTAACCGGTTTAGACCGTACCGGCTTCC	29560
QY	2641	tggagatctgtctgtagacttaccaaaattatagcaagaatgggtctattatatttaa	27000
Db	29562	TGGGATTTCTCTGTGACGATTTATCCAAAATATATAGCACGAATGGGTTATTATTTAA	29500
QY	2701	aactcacaacttgaatcagataaaaattcatalaaacactttaaagatgctcgtagatct	27600
Db	29502	AACTCACAACCTTGATGATGAATTAATTTCAATTAACACTTTTACGATGCAATGCTGATGATCT	29440
QY	2761	atcctaagacttttttttctcacaacggtagatgaagaatrtatgatatgacagag	28200
Db	29442	ATCTAATGACTTTTTTTTTTCTACCAACGGGTGAGTAAGAACTTATATGCTATTATGACCAAG	29380
QY	2821	acaattgattatgatatataccaattacacatgatatcttattgatatataatgctgtttaa	28800

```
|||||
Db 29382 ACATTGATATATGATATATCCATATATCATGATATATGATATATTAATAGCTGTAA 29323
|||
Qy 2881 ctattcagcatcgcaagcttctcgcaacttctgttcttaactttaagaagtttaataa 2940
|||
Db 29322 CTATTCAGATGCGACCTTCTGCAACTTTGTTTTAATTTAAGAGTTAATAATAA 29263
|||
Qy 2941 aatattaaagaagcatcagagcaaaaagtaatgaacacgagaaacaaaagca 3000
|||
Db 29262 AAGTATTAAGAGCATACGAGCAAAAGTAAATGACACGAGAAACAAAGCA 29203
|||
Qy 3001 tgaagcatctggtta-gtttaagcttaataagaagaattatbaattttaagcat 3059
|||
Db 29202 TGAAGCTCATTTGGTTAGCTTTAAGCTTAATAGAATTTTAAATTAATAGAGAT 29143
|||
Qy 3060 gataacaattatattcttcgaactctttaaaacccctcttaacacagaagctccctc 3119
|||
Db 29142 GATTAACAAATTAATTTCTGACTTCTTAAACCCCTCTTACAAACAGAACTCCCTTT 29083
|||
Qy 3120 ttcaagtagagtcgacttcccaatccttaagaacaaagccattagaagagaagtgatg 3179
|||
Db 29082 TTCAAGTAGAAGCTCGATTTCCCAATCTTAAAGACAAAGCCATTAGAAAGAGAAAGTG 29023
|||
Qy 3180 agagagagagagaactagctccatgctcgaacagagacatcaactctcttaactctc 3239
|||
Db 29022 AGAGAGAGAGAGAAACTAGCTCCATGTTGAAACAGAGACATCATCTCTTACCTCTTC 28963
|||
Qy 3240 ttctctcccatcgcttctgtctctctctctctctctctctctctctctctctctct 3299
|||
Db 28962 TTCTTCTCCCATGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 28903
|||
Qy 3300 gaaaaacagatcactcactcctcgggttaactcgggttgagcaattctctgtgtaaacca 3359
|||
Db 28902 GAAAAACAGATTCATCTTACTCTCGGGTAATCCGGTTGGCCCTTTCTTGSTAAACCA 28843
|||
Qy 3360 tccggtatcttaaacggtlacacgcacaaactcgggtgaactctgaactctgaacaaact 3419
|||
Db 28842 TCGGTATCTTTAAACCGGTACACCGCACAAACACTCGGTGACTTCATGCAACAACTG 28783
|||
Qy 3420 ccaagtaaaacaaacactctccaaaaactcaaaaaataactctctgttcttgaatt 3479
|||
Db 28782 CCAAGTAAACAAACAACTCTCCAAAACCTCAAAAATAATCTCTGTTTGAATTT 28723
|||
Qy 3480 tgaactatgtctgttattcttaacagatatggtlaagatatataagctgaactgtttgaga 3539
|||
Db 28722 TGACTATATGTTGTTATTTTACAGAGTATGGAAGATATATAGATCGAACTTTTGGA 28663
|||
Qy 3540 accaagcatcgtatcagctgactgctgaacttaataatcaatatacaaaaaggaag 3599
|||
Db 28662 ACCAAGCATGCTATCAGCTGATCGTGACCTTAATAGATTTCATTTACAAAACGAAG 28603
|||
Qy 3600 gctcttgaatgtatgtatctcctaagaatataagtggaactcttggaatgtctgact 3659
|||
Db 28602 GCTCTTGAATGTATGTTATCTTAAGATATAGTGGAATCTTGGAAATGCTGATGCT 28543
|||
Qy 3660 tgtctgttctgtgtagatgcatagatagatagagaatcgcgttaactcttaactca 3719
|||
Db 28542 TGTTCTTGTGTGTCATAGCATAGATATGAGATATGCGGTTAACTTTAAAGTCA 28483
|||
Qy 3720 cgaagctcttagaactatctacttaagaatgttagagagacatacttctgttctgta 3779
|||
Db 28482 CGCAGCTCTTAGAATCTTACTTAAAGATGTGAGAGACATCTTGTGTTCTTGA 28423
|||
Qy 3780 ttcttgcaacaaactcttctctctctctcgaagacgaagccaaaagtttatttt 3839
|||
Db 28422 TTCTTGCAACAAACCTTATTTCTCTGCTCAAGACGAGGCCAAAAGGTTTTATTTT 28363
|||
Qy 3840 tatcttattctgctaatttttctgttcttaagacttgaagttcttaactttttt 3899
|||
Db 28362 TATCTTTTATTTTCTTAAATTTTCTTGTATGATCTTTTGAATTTCTTAACTTTT 28303
|||
Qy 3900 tttaactgaacagtttaagtttaactaaatgacgaagacataatgaatagatgactctg 3959
|||
```

QY 5039 atctcagcggtacatttgatataatctcgtltatgaccaactaactctttaaactcttg 5098  
|||||  
Db 27162 AACTCAGCCGTCATTTGGATTAATCTCGTTATGACCAACCTAATCTTTAATCCTGG 27103  
QY 5099 agatgcaacaggtataataaagttctctcgttactatcgaaataatagttagtt 5158  
|||||  
Db 27102 AGATTGCAACAGGTAAATAAAGTTCTCTCGTTACTATGCAAAATTAAGTGTATAGTT 27043  
QY 5159 ttctcatctatctcatgaaatagatacgtctcgttacttactatctatagatactata 5218  
|||||  
Db 27042 TTTTCAATCTATTGATGATAGTATAGTCTCTAGCTGATTTACTATCTATAGTATATA 26983  
QY 5219 cgaagaactatcaatctgcaaaaaacttctatatactatcttcaagtlagacttaa 5278  
|||||  
Db 26982 CGAGACATTTAATCTGCAAAAACTTTTATTTATTTATCTTTCAGTTAGATCTTAA 26923  
QY 5279 caagtcatactatctatgatacacaatgaaacataataataaataagagaagaag 5338  
|||||  
Db 26922 CACGTATGGATCATTTGATCAGATGAAAGCATTAATAAATAAAGAGAGAAAGAG 26863  
QY 5339 acgtgtgtgtatgatactgatacgttgaagaacataatagttagttatgcttctta 5398  
|||||  
Db 26862 ACCTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26803  
QY 5399 cgttagagctgctcaataatctctataatcgttgaacgttgaattatctatagtcag 5458  
|||||  
Db 26802 CGTAGAGCTGCTTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 26743  
QY 5459 ttgatacaatttaagaatttgagaagaagatgatacgaacgttgaagaagatctcaac 5518  
|||||  
Db 26742 TTGATTAATTTAATTAAGATTTGACGAGATGATGATGATGATGATGATGATGATGAT 26683  
QY 5519 cgaatgctgtatgatactgatactctgcaacgttgaacgttgaattatctatagttat 5578  
|||||  
Db 26682 CAGTATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26623  
QY 5579 aactctatataatacaactatcaaatcgcgaatlagagaataatactataatagttat 5638  
|||||  
Db 26622 AACTCATTTAATTTAATTAAGATTTGACGAGATGATGATGATGATGATGATGATGAT 26563  
QY 5639 tatatacatatataatgataatatacttgaataaagaataaagaatgataatagttat 5698  
|||||  
Db 26562 TATTAATGATTTAATTAAGATTTGACGAGATGATGATGATGATGATGATGATGATGAT 26503  
QY 5699 tataagtgatgagaagaagtgatgataatgataatgataatgataatgataatgataat 5758  
|||||  
Db 26502 TATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 26444  
QY 5759 tgcgaacaaagcagtgct 5818  
|||||  
Db 26443 TGCACAACGACAGTGTCT 26385  
QY 5819 ttctgataataaatacgttctgcttctctctctctctctctctctctctctctctctct 5878  
|||||  
Db 26384 TTGTGATCTATTAATAATCGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26325  
QY 5879 gtcatactgataagaagtgatgataatctctgacgttggagaacaaactatactgacgt 5938  
|||||  
Db 26324 GTGATCTCTCAGGAAGTGTGATTTTTCGAGTGGGCAACAACTACATGCGCTTGGAGG 26265  
QY 5939 agggccaagcctatgtctgctgctcgaagctagcgaagttagaataatgcaagttctat 5998  
|||||  
Db 26264 AGGGCCAAGCCTATGTGCTGCTCAGAGCTAGCAAGTTAGAAATGGCAGTGTATTTCA 26205  
QY 5999 tcatctagttcttaataatcaatgggaatlagagaagaatgatacaacattctctctct 6058  
|||||  
Db 26204 TCATCTAGTTCTTAATTAATTTGGAATTTAGCAGAAATGATTAATCCATTTGCTTTTCC 26145  
QY 6059 ttctgttgaattctctcaacggttgcctatagaggttctctgctatctglaaaaaaaa 6118  
|||||  
Db 26144 TTTTGTGATTTTCTTAACAGGTTTGGCTTATAGGTTTCTCGATTTCTGTAAAAA 26085

QY 6119 aaagatgaagaatcttct 6178  
|||||  
Db 26084 AAGATGAAGAAGTATTTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 26025  
QY 6179 gcccaatgataataaatttgataaataatattatgattgattcgttcttctgctg 6238  
|||||  
Db 26024 GCCCAATGATATTAATAAATTGGATTAATTAATTAATTAATTAATTAATTAATTAAT 25965  
QY 6239 gtttgaagaagaaggttctgacttctgaaatgagcagatgatalatagattggaagctag 6298  
|||||  
Db 25964 GGTTTAGAAAAGGCTTTGAGCTTTGCAAGAGTGACGATGATATGATTTGGAGCTTAGG 25905  
QY 6299 tgaagtccttggacatttgatgagtgatgagtgatgagtgatgagtgatgagtgat 6358  
|||||  
Db 25904 TTGAGCTTTTGGACATTTTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 25845  
QY 6359 ttaaatggccttctataagggcccaattatagattatatacaagaatgaaacttcta 6418  
|||||  
Db 25844 TTAAATGGCTTTCTATTAAGGCCCAATTAATTAATTAATTAATTAATTAATTAATTAAT 25785  
QY 6419 ctctgcttctgatacgaagaacataaacttgcataatacacaagaattatgata 6478  
|||||  
Db 25784 CCGTCTTTTGAATCCGACAAATTAACAAATTTGCAATATGCAAAACGACAAATTAATGTA 25725  
QY 6479 aaactgctgtgctatgattggaataatcattggcgtgagacgtgatacaagaacaga 6538  
|||||  
Db 25724 ACACCTGTGTGCTGTAGTGGGAAATCATTTGGCTGGAACATGACATGACAAACAA 25665  
QY 6539 aaactgcaatataagataacactcctatgagaaggttccaaactatcttactatct 6598  
|||||  
Db 25664 AACCTGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 25605  
QY 6599 cgttatacatctgacaagaagtgatgatacgaagaagcctaaatgatacaataatgctc 6658  
|||||  
Db 25604 CGGTGTACATTTGACAAAGAGTGTAGTGTCAAAAGAGCTAATGATTAACATTAATCTGT 25545  
QY 6659 caagataagaaggttcaataaattgcttctctctctctctctctctctctctctctct 6718  
|||||  
Db 25544 CAGGATTAAGAGGCTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 25485  
QY 6719 tactgttcttgggct 6778  
|||||  
Db 25484 TACTGTCTTGGGCTCTATTTCTCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 25425  
QY 6779 catcatctttagatct 6837  
|||||  
Db 25424 CATCATCTTTAGATCT 25365  
QY 6838 cctcatcttctgaaagaca-gtctcttccaaagtcctagtcagggccaaagctc 6888  
|||||  
Db 25364 CTCCTCTTTTGAAGACAGGCTCTTTTCCAAAGTCTGTAGTCAAGGCCAAAGCTT 25313  
RESULT 2  
AF044216 4818 bp DNA PLN 06-MAR-1998  
LOCUS Arabidopsis thaliana steroid 22-alpha-hydroxylase (DWF4) gene,  
DEFINITION complete cds.  
ACCESSION AF044216  
VERSION AF044216.1 GI:2935341  
KEYWORDS  
SOURCE  
ORGANISM  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
Brassicales; Brassicaceae; Arabidops.  
REFERENCE  
1 (bases 1 to 4818)  
Choe, S., Dilkes, B.P., Fujioaka, S., Takatsuto, S., Sakurai, A. and  
Feldmann, K.A.  
The DWF4 gene of Arabidopsis encodes a cytochrome P450 that  
mediates multiple 22 alpha hydroxylation steps in brassinosteroid  
biosynthesis  
JOURNAL Plant Cell 10 (1998) In press

REFERENCE 2 (bases 1 to 4818)  
 AUTHORS Choe, S., Dikes, B. P., Azpiroz, R. and Feldmann, K. A.  
 TITLE Direct Submission  
 JOURNAL Submitted (22-JAN-1998) Plant Sciences, University of Arizona,  
 Tucson, AZ 85721, USA

FEATURES  
 source  
 Location/Qualifiers  
 1. 4818  
 /organism="Arabidopsis thaliana"  
 /strain="Ws-2"  
 /db\_xref="taxon:3702"  
 /chromosome="III"  
 /map="8 cm from nga6"  
 990..4398  
 /gene="DWF4"  
 /note="dwarf4"  
 /allele="dwf4-1; dwf4-2; dwf4-3; dwf4-4"  
 990..1055  
 /gene="DWF4"  
 join(<1133..1353,1434..1758,1844..1996,2095..2409,  
 2562..2654,2746..2824,2931..3040,3795..4398)  
 /gene="DWF4"  
 /product="steroid 22-alpha-hydroxylase"  
 join(1133..1353,1434..1758,1844..1996,2095..2409,  
 2562..2654,2746..2824,2931..3040,3795..4040)  
 /gene="DWF4"  
 /function="brassinosteroid biosynthetic enzyme"  
 /note="member of the cytochrome P450 superfamily; CYP90B1"  
 /codon\_start=1  
 /product="steroid 22-alpha-hydroxylase"  
 /protein\_id="AAC05093.1"  
 /db\_xref="GI:2935342"  
 /translation="MFEENHTLLPLLLPSLLSLFLFLKRRNRKTRFNLPPGKS  
 GMPFETIGLYKPYATTLGDEPMOHSYKGYIRSNLGEPTIVSADGLKRFILQ  
 NGRLEPCSYRSTIGLIGKSMWLYVGMHMRSTSLNLSHALRFLIKLVERH  
 TLEFVDSMOONSIFTSKODEKKTFLMLAKHISMORGEETQLKREYTERKGVVS  
 APLNLPSTAYHKAQSRATLTKFERKMERKLDIEEDDEEVEVTEDEAKSKSDH  
 VRKQRTDDLGVNLKHSNLSQIDLLISLFAQHSVAIALAIFLOKCPKAV  
 EELREHLEIARAKKEIGSELNMDYKKMFQCYINETRLGNVRFARAKLADY  
 RYKGYDIPSGMKLPIVSAVHLNDRYDOPNLEPMRMOONNGASSGSGSTWGN  
 NYMPGGGPRLCGSELAKLEMAVFTHHLVTKRWELAEEDDFAPFPVDFPGLPIR  
 VSKRL"

BASE COUNT 1566 a 727 c 888 g 1637 t  
 ORIGIN

Query Match 69.9% Score 4818; DB 12; Length 4818;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 4818; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2071 gaattcaatatccatataatcttgaactcaaatataatattagccaat 2130  
 |||||||  
 Db 1 GAATTCAAATATCTCCATTAATATTTTGAACATCTACAAATTTATATATAGTCAAT 60  
 |||||||

QY 2131 acaaatgcatagaaagttccaaaaaaatttgttaacagaactccaaattttttt 2190  
 |||||||  
 Db 61 AACAAATCATAGAAAGTTCCAAAAAATTTGTTAACAGAACTTCCAAATTTTTTTT 120  
 |||||||

QY 2191 ttcatggaacagaataacagatagaactatttctgttggtaacggaagtata 2250  
 |||||||  
 Db 121 TTTATGGAACAAGAAATACAGATGAAACTATTTTGTGTGGAATGGAAGTGTATATA 180  
 |||||||

QY 2251 tacattgaacaaattttaaaaaattatataagcctatacgcgcccaagtatgtatcta 2310  
 |||||||  
 Db 181 TACATTAGCAAAATTTTAAAAAATTTATATAGCCTATACGCGCTCAAGATGTATTATCTA 240  
 |||||||

QY 2311 gtaggtgtaataaataatgcatggctcagaattgggacacaacatgaacacggaat 2370  
 |||||||  
 Db 241 GTAGGTGTAATTAATATGATGATGCGAATTCAGAAATGGGACAACATGAAACGGGAAT 300  
 |||||||

QY 2371 taaaataataacttaataataataaatttgagtaaatgtgtcttcctgactattgag 2430  
 |||||||  
 Db 301 TAAAAATTTAACTTTAAATAAATTTGATGAAATGTGTTTCTGACTATGAGG 360  
 |||||||

QY 2431 gcaaaaaaagacatcgcgaagaagctacaggttctgactgtccagttcgttaataact 2490  
 |||||||  
 Db 361 GCGAAAAAAGACAAATGCGCCAAAGTCTACGGGTTTGACTCTCCAGTTGGTATTAATCT 420  
 |||||||

QY 2491 aataactgtcttctgacgcgacgtcgtgtagggtctctctgacatttctactgtctc 2550  
 |||||||  
 Db 421 AATAACTGTCTTGTGACCGCACGCTCGTGAAGGGTCTTCTGACATTTTCACTGTTCCT 480  
 |||||||

QY 2551 accctctcgttgagcccaacctttcccatatctctaaggtaattttggaaccccaat 2610  
 |||||||  
 Db 481 ACCCTCTCTGTGAGCCACCCCTTTTCCATATCTTAAGGTAATTTGGAAATCCCAAT 540  
 |||||||

QY 2611 ttaaccgatgtgaacgcgttacccgacttccctggatctctgttgagacattatcaaaat 2670  
 |||||||  
 Db 541 TTAACCGATTGACCGGTACCGGACCTTCCGGGATTCGTGTGAGCATTTATCAAAAAT 600  
 |||||||

QY 2671 tattaagaagaatggttatttaataaaactcaacacttgatcagataaaattcat 2730  
 |||||||  
 Db 601 TATTAGACGAAATGGGTTTATTATTAATAAAACTCACAACTTGATGATTAATTTTCAT 660  
 |||||||

QY 2731 aaacactttacagatgattcgtacgactctatctcaatgacttttttttctaccaggt 2790  
 |||||||  
 Db 661 AAACACTTTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 |||||||

QY 2791 ggaatgaagtatatagtaactatacgaagagaacattgataatataatcattatcc 2850  
 |||||||  
 Db 721 GGATGAAGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 |||||||

QY 2851 atgatatattatgataataatagctgttaactatttcagacatgcagactttctgcaact 2910  
 |||||||  
 Db 781 ATGATATTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840  
 |||||||

QY 2911 ttgtttttaaatttaaggttttaataaaagratataaaaggagcctaaagagagaca 2970  
 |||||||  
 Db 841 TTGTTTAAATTAAGGTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900  
 |||||||

QY 2971 aagaatgaaacagaggaacaaagccatgaagctcaatggttatttgaagcttaata 3030  
 |||||||  
 Db 901 AAATTAATGAACACAGGAACAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 960  
 |||||||

QY 3031 agaaatttatataatttaataatgaagatgaatacaattatattcttgactctttaa 3090  
 |||||||  
 Db 961 AGAAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020  
 |||||||

QY 3091 accccctcttaacaaagaaactcccttttcaatagaagtcgaatcccaatttaag 3150  
 |||||||  
 Db 1021 ACCCCCTCTTAACAACGAAGCTCCCTTTTCAAGTAGAAGTCCGATCCCAATCTTAAG 1080  
 |||||||

QY 3151 acaaaagccattagaagaagaagtgagtgagaagagagaagaagaactagctccatgctcga 3210  
 |||||||  
 Db 1081 ACAAGGCATTTAGAAAGGAAGGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
 |||||||

QY 3211 aacagagcatcaatactctcttaactctctctctctccatcgctttgtctctctctc 3270  
 |||||||  
 Db 1141 AACAGAGCATATATCTCTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200  
 |||||||

QY 3271 ctctctgattctctctgaaagaagaagaataagaacaaacgaattcaatctactccgggttaa 3330  
 |||||||  
 Db 1201 CTCTCTGATTTCTCTTGAAGAAAGAAATGAAAAACCGATTCAATCTCTCTCTCTCTCT 1260  
 |||||||

QY 3331 atccggttgagcattctctgttgaaacatcggttatttaaacccgtacacgcgcacac 3390  
 |||||||  
 Db 1261 ATCCGGTGGGCAATTTCTGTGTGAACCATGCGTTATCTTAACCGGTACACCGCACAC 1320  
 |||||||

QY 3391 actcgtgacttcatgcaacacacatgctccaaglaaacacacacatctccaaaactc 3450  
 |||||||  
 Db 1321 ACTCGTGACTTCAATGCAACACATGCTCTCCAGTAACACACATCTCTCCAAAACCTC 1380  
 |||||||

QY 3451 aaaaaataaactcctcgtcttggaaatttgaccaaagtgttcaatttaaggataggt 3510  
 |||||||  
 Db 1381 AAAAAAATAAATCTCTGTTTGTGAAATTTGACATTAAGTTGTTTATTTTACAGGTAATG 1440  
 |||||||

QY 3511 aagatatatagatcgaactgttctgggaacacacgatcgatcagctgactgagactt 3570  
 |||||||

```
|||||
Db 1441 AAGATATATAGATCGAAGCTTTGGAGAACCAACGATCGTATCAGCTGATGCTGACCTT 1500
QY 3571 aatagattcatatatacaaaacgaagaagcctttgaatgtagtattatcctagaatata 3630
Db 1501 AATAGATATATATATACAAAACGAAGAGGCTTTGAAATGATATCTTAAGATATA 1560
QY 3631 gttggaattcttggaatatgtagcagctgtagtctgtagttagacatgcatagaatag 3690
Db 1561 GGTGGATCTCTTGGAATAATGATGATGCTTGTCTGTGTGTACATGATTAAGATATG 1620
QY 3691 agaagatctgcgttaactcttaagtaacgacgctcttaagaactatctactaaagat 3750
Db 1621 AGAAGATATCTGCTTAACCTTTAACTCAACGACGCTTAGAATCTGCTACTTAAGAT 1680
QY 3751 gttgagaagacatacttctgtagtctgtagtcttggcaaaaacactatcttctgct 3810
Db 1681 GTTGAAGACATATCTTTGTTGTTGTTGATCTTGGCAACAAACCTATTTTCTGTGCT 1740
QY 3811 caagaacgaagccaaaaggctttatcttattatcttattcttgtaaatcttctgta 3870
Db 1741 CAAGACGAGGCCAAAAGCTTTTATCTTTATCTTTATCTTTGCTAAATTTTGTGTTA 1800
QY 3871 tgaattcttaagagttcttaacttttttttttaattgaacagcttcaagtttaactaat 3930
Db 1801 TGAATCTTAAAGATTTCTAACTTTTAACTTTTAACTTAACTTAACTTAACTTAACT 1860
QY 3931 ggcgaagacataatagatagatagatcctgagaagaagaacagagaacttaagaaga 3990
Db 1861 GCGAGAGCATATATATAGATATGATGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 3991 gtagtaacttcaatgaagaagtagtctgtagtctgtagtcttaactcaacgaagctcta 4050
Db 1921 GTATGTAACTTTTCAATAAAGAGATGCTGCTGCTTAAATCTTAACTTAACTTAACTTAA 1980
QY 4051 tcaataagctctcaagtagcatatttttttggtagtaagtaacactctctat 4110
Db 1981 TCAATAAAGCTTTCAGGATCAATTAATTTTGTGTAAGTCAACAACTCTCAATTAAT 2040
QY 4111 aggttttaatttatttattgtagttaaatactcaaaatgtagttagtaacga 4170
Db 2041 AGGTTTTTAAATTTTATTTATGTTAAATTAATTAATTAATTAATTAATTAATTAATTA 2100
QY 4171 gcaagaatttgaaatcttgagagagaatgaagaagaatctgaatcaagaaga 4230
Db 2101 GCAACATATATGAAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 4231 gaagatcaagaagaagaagaatgaagaagaagaatgaagaagaatgaagaagaatgaaga 4290
Db 2161 GAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 4291 catgttaggaagaagaagaagaatgaagaagaatgaagaagaatgaagaagaatgaaga 4350
Db 2221 CATGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 4351 ttatgaagcgaagaagaatctgtagtcttcttgaatgtagttagttagttagttagttag 4410
Db 2281 TTAATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
QY 4411 tcttcgttagcattgtagtctgtagtcttcttcttgaagcgttgaagcgttgaagcgttga 4470
Db 2341 TCTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
QY 4471 gaagcttaggtaagaatatacaagaagaatgaagaagaatgaagaagaatgaagaagaat 4530
Db 2401 GAGCTTAGGTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2460
QY 4531 atataagatataagaatataatcttgaatatacagtagtaagaagaagaatgaagaagaat 4590
Db 2461 AATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2520
QY 4591 atctgcaactatctttagttagttagttagttagttagttagttagttagttagttagttag 4650
|||||

Db 2521 ATTGTCATTAATTTATGTTTATGATTTATGATTTTGAAGAGAGAGATCTTGAGATCG 2580
QY 4651 cgaagggccaaagaagaagaatcgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 4710
Db 2581 CGAGGGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
QY 4711 accttaacatgtagttagttagttagttagttagttagttagttagttagttagttagttagt 4770
Db 2641 ACTTACTCAATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 4771 atgaagaaccaaataatgaatcttcttcttcttcttcttcttcttcttcttcttcttctt 4830
Db 2701 ATGATTAACCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2760
QY 4831 ctgcgaatggaagaatgaagaatgtagttagttagttagttagttagttagttagttagttag 4890
Db 2761 CTTCGATGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
QY 4891 aaagtaaaactttagttagttagttagttagttagttagttagttagttagttagttagttagt 4950
Db 2821 AAAGTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2880
QY 4951 ttggaagaagaatataaataatcttcttcttcttcttcttcttcttcttcttcttcttctt 5010
Db 2881 TTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
QY 5011 cccatgtaggtagaagtagttagttagttagttagttagttagttagttagttagttagttagttagt 5070
Db 2941 CCTTACTGGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
QY 5071 tgaacaaacttaacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 5130
Db 3001 TGACCAACCTTAATCTTAACTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
QY 5131 gtttaactcgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaagaatgaaga 5190
Db 3061 GTTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
QY 5191 cgtgatttaacttcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 5250
Db 3121 CGTATTTACCTATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
QY 5251 tattattatcttcaagttgaatcttcaacagctgaatgaatgaatgaatgaatgaatgaatgaat 5310
Db 3181 TATTATTATCTTCAAGTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
QY 5311 ataaatataaataaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 5370
Db 3241 AATAATTAATAATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
QY 5371 aattagtagtagttagttagttagttagttagttagttagttagttagttagttagttagttagt 5430
Db 3301 AATTAGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
QY 5431 gaccgttagttagttagttagttagttagttagttagttagttagttagttagttagttagttagt 5490
Db 3361 GACCGTTGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3420
QY 5491 accaagcctttagcgaatctcaacacgcccgaatgaatgaatgaatgaatgaatgaatgaatgaat 5550
Db 3421 ACCAGGCTTTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3480
QY 5551 gttgactggttttaatacttagttagttagttagttagttagttagttagttagttagttagttagt 5610
Db 3481 GTTGACTGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3540
QY 5611 cgaatagaagaataaataaataaataaataaataaataaataaataaataaataaataaataaataa 5670
Db 3541 CGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
QY 5671 aataagaataaataaataaataaataaataaataaataaataaataaataaataaataaataaataa 5730
Db 3601 AATAAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
|||||
```



[illegible]

QY	6811	aaataactctgtccacgaacatggtctcaatctttgaagcaagctctttccaagtc	6870
Db	4741	AAGTATATCTCTGTTACCAAGCAATGGTCTCATCTTTGACAGACAGGTCTTTTCCAAGTCC	4800
QY	6871	tagttcagcgcaaacgctt	6888
Db	4801	TAGTTCAGGCCAAAGCTT	4818
RESULT	3		
ATF18022			
LOCUS	Arabidopsis thaliana	DNA	PLN 13-APR-2000
DEFINITION	Arabidopsis thaliana DNA chromosome 5, BAC clone F18022 (ESSA project).		
ACCESSION	AL163817		
VERSION	AL163817.1	GI:7573446	
KEYWORDS			
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota, Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.		
AUTHORS	1 (bases 1 to 100269) Bevan, M., Murphy, G., Ridley, P., Hudson, S., Bancroft, I., Mewes, H. W., Rudd, S., Lemcke, K. and Mayer, K. F. X.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 100269) EU Arabidopsis sequencing project.		
AUTHORS	Direct Submission		
TITLE	Submitted (13-APR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG. E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK.		
JOURNAL	E-mail: michael.bevan@bsrc.ac.uk		
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <a href="http://www.mips.biochem.mpg.de/proj/thal/">http://www.mips.biochem.mpg.de/proj/thal/</a> .		
FEATURES	Location/Qualifiers		
source	1..100269		
	/organism="Arabidopsis thaliana"		
	/variety="Columbia"		
	/db_xref="taxon:3702"		
	/chromosome="5"		
	2255..6118		
	/gene="F18022_10"		
	complement(join(2255..2376,2496..2602,2699..2773,2896..3004,3094..3217,3419..3493,3582..3743,3871..4005,4364..4393,4488..4553,4723..4759,4849..4922,5124..5180,5289..5339,5466..5531,5640..5775,5930..5981,6091..6118))		
	/gene="F18022_10"		
	complement(2255..2376)		
	/gene="F18022_10"		
	/number=1		
	complement(join(2255..2376,2496..2602,2699..2773,2896..3004,3094..3217,3419..3493,3582..3743,3871..4005,4364..4393,4488..4553,4723..4759,4849..4922,5124..5180,5289..5339,5466..5531,5640..5775,5930..5981,6091..6118))		
	/gene="F18022_10"		
	/note="strong similarity to protoporphyrinogen IX oxidase, glycine max, AB025102"		
	/codon_start=1		
	/product="protoporphyrinogen oxidase-like protein"		
	/protein_id="CAB87761.1"		
	/db_xref="GI:7573447"		
	/translation="MASCAYADHQIEAVSGKRVAVYAGVSGLAGAARVILKSRGLNVTVEADGVRGKLRISWQNLIMDEGANTITEAPEVGSLLDGLERKQFPISOKRYIVRNGVPMPLPTNIELVTSSVLSTQSFQILLEPFLMKRSKVSADSAEVSSEFKRSRHKGEVADILIDPEVGTGAADPDSLSMKHSPDLNWSGSIIVGAIPIKFAAGKQSHQSPGTRKSGRSFSGKQMLDPTFLCKLSHDEILNDSKVLSTLNSGSRQENMLSTSHNETRONHYNDAAPLCNLYKEMKVYKGGQPOLNPLPEINWPMLSYLITTFTEAKRRLPEGRFVILPSKEQKHGFTLTGLTSSMKFPDRPSDVLITTFTEGGRNOELAKASTDELKQVNTSDIQLRLGVEEEPVSVMHYWRKAFPLYDSSYDSVMEADIK		
CDS			
exon			
gene			
.gene			







[illegible]

QY	2061	aatagctggcgatccaatccatcaatc----	attcttggaaacctcaaatatt	2116
Db	74423	TATATTGCAATATATTTTATTTTAAATTAA	TATATATATATATATTTATTTATTTAT	74482
QY	2117	aatacttcgtcaatacaacatgcatacgaa	agctcccaaaaaaatttctgtacagaaact	2176
Db	74483	TATTTTAAATTAATTAATTAATTAATTAAT	TATTAATTAATTAATTAATTAATTAAT	74542
QY	2177	ccaactcttcttcttctatgacacaagaata	catacagatacgaaactattctctgga	2236
Db	74543	ATTATTTATTTTATTTATTTTAAATTAAT	TATTAATTAATTAATTAATTAATTAAT	74602
QY	2237	tggagctgtaataatacattgaagcaaatct	ttcaaaaaattataaagccatacgcgtca	2296
Db	74603	TATATTTATTAATTAATTAATTAATTAAT	TATTTTATTTTATTTTATTTTAAATTAAT	74662
QY	2297	aagtactgatactcagtcagtgtaactaaat	gcatcgcagatcgaggacaac	2356
Db	74663	ATTATTAATTAATTAATTAATTAATTAAT	TATTAATTAATTAATTAATTAATTTATTTT	74720
QY	2357	aatgaaacgcgaatataaacttaaccttaaa	ataataaataaacttga	2404
Db	74721	ATTTTAAATTAATTAATTAATTAATTAAT	TAATTAATTAATTAATTAATTAATTAAT	74768

RESULT	5
LOCUS	AC004157
DEFINITION	AC004157 169546 bp DNA HTG 12-AUG-2000 Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.
ACCESSION	AC004157
VERSION	AC004157.8 GI:9797712
KEYWORDS	HTG; HTGS_PHASE1. malaria parasite P. falciparum.
SOURCE ORGANISM	Plasmodium falciparum Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE AUTHORS	1 (bases 1 to 169546) Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,T., Tamaki,T., Kirdl,O.B., Conway,A.B. and Davis,R.W. Plasmodium falciparum 3D7 chromosome 12 Unpublished
JOURNAL TITLE	2 (bases 1 to 169546) Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W. Direct Submission
JOURNAL	Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	On Aug 12, 2000 this sequence version replaced gi:8810447.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

* 1	23466: contig of 23466 bp in length
* 23467	23666: gap of unknown length
* 23667	169546: contig of 145860 bp in length.

```

FEATURES
  source      Location/Qualifiers
1. 169546    /organism="Plasmodium falciparum"
           /db_xref="taxon:5833"
           /chromosome="12"
           /clone="PFYAC293"
           /clone="3D7"
BASE COUNT   69871 a 15381 c 15705 g 68389 t 200 others
ORIGIN

```

Query Match 1.9%; Score 133.2; DB 60; Length 169546;  
 Best Local Similarity 44.7%; Pred. No. 2.1e+09;  
 Matches 1041; Conservative 0; Mismatches 1248; Indels 39; Gaps 12;







D	85844	TTATATACATTAATTAATATATATATTTTATTTTAAAAAAATTAATTAATTAATTAAT					85785
Q	639	gagtagtgacccaatggagtagtgcgaataatctcaagaagataagagatgcaaccaa					698
D	85784	TATATAATTTAAATATATTTTAATAGTAAATTTAAATTTAAACAAATTAATTAATTAATA					85725
Q	699	aggcttggaaataatgctcccgccagctctctccacatacatatcgaccctattgg					758
D	85724	AATATTAATTAATTAATTAATAAATAAATTAATTTTATATATATATATTAATAT					85665
Q	759	attcttcggaatctcgttaaaatctgcgaatacagatlglaaaaatatttatctgtag					818
D	85664	TAAATTTAAATTTAAATTTAAGA-----AAGAAATATTTTAACTTTTA					85623
Q	819	ctgactcgaataatagctcccgagatctgcgaatactctctcttaagcatattctgc					878
D	85622	TTTAAATTAATTAATTAAGTAAATTAATTTTATGTTATT-TATTAATTAATTAATTAATA					85564
Q	879	tctcttctgtctcgtctctcctaataatactcgcgaataatgataaataatgata					938
D	85563	TTTTATTTTATTTATTTAATTAATTAATTAATTTTAAATATATTTTATTTAAATTTAA					85504
Q	939	tatcacaacaacatgctcgggacacatttgataaaccttctccaaacattcaggac					998
D	85503	ATATAAACCTAATTAATTAATTAATTAATTAATTAATTTTAAATTAATTAATTAAT					85444
Q	999	actggactcgacccttaaaatacgaattctacgcgcactagctagatctacgacata					1058
D	85443	ATATTTTACATTTTATTAATTAATAAATAAACCATTTATTAATTAACCTTAATTAATAT					85364
Q	1059	agcataaaggaccggttcaagctatttacaagaatllacaacatgaatatagcttgaat					1118
D	85383	AAAAATTAATTAATTAATTTGTGTAATTTAAATTTAAATTAATAACATTTTAAATTAAT					85324
Q	1119	cccttagaanaatctggaaatcacgctgtctatglaaataatagatttagtgglaacaa					1178
D	85323	TATATATATATTAATTAATTAATTAATTAATTAATTAATTTTAAATTAATTAATTAAT					85264
Q	1179	ta-----tgttaacaaatcagtggtcacaataacataacaaatccctcagaaaaacaa					1234
D	85263	AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAAATTAATTTTAA					85204
Q	1235	taaggaagatcaacatccatacatatggtatgcataaccttcacgctatgctatacag					1294
D	85203	TAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAAATTAAT					85144
Q	1295	agactaaagaa-----tagtataatgctatgcataaataatgaatcacacgctgtat					1348
D	85143	ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT					85084
Q	1349	aattctggagacgctatgtagcaatcaacgcaaatcaatctctctgtgtgcacaataa					1408
D	85083	TTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT					85024
Q	1409	aacaaacaaagaaaaaagaagaacgatttctctggaatccatccaatgactcaaatg					1468
D	85023	ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA					84964
Q	1469	catgatactttgggttacagttccgaagctcctcagaacgctglaaacacatctgcaact					1528
D	84963	TTTTAATTAATTTTAAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT					84904
Q	1529	ttaaatctgctcttcaatgcaatcttcaacatattatctgttagcttgaatllaaga					1588
D	84903	ATATATATTTAATTTAATTAATTAATTAATTAATTAATTAATTTGATATTAATTTTAAT					84864
Q	1589	gcgaactctgtaacatcacataatctataatgatactagatctgatatctatccaataca					1648
D	84843	TATATATATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTTAAAC					84784
Q	1649	tacttctgagtgcttaaccttaactctgttctctcctaaggtataaataatcatctgag					1708

Dd	84783	TATTAATTATTTAATTAAATTATTAATTATTTATTTTATTTACTAAATAATTAATTAATTAT	84724
Qy	1709	gcaaaaaaaagttctgtcattatcttcgcgatgcatagaagaataaacctaagacttaac	1768
Dd	84723	AATATATTTATGTTTTAAATTAATTTAAATCAATTCATTTTATTTAAATTAATTTATATTTAT	84664
Qy	1769	tcttgaanaatgtaacccttctacctaatcatgaatcaatcacagtctgttcttgccataat	1828
Dd	84663	ATTAAATATATCTTTAATTATTAATAATAATTAATGTTTAATATATATTAATTAATTAATA	84604
Qy	1829	gacagcccttacacactgtagtccaattcttcctgcacaatatcaaatlagaalccaat	1888
Dd	84603	TTTAAATATATGTTAATTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84549
Qy	1889	gctactactcaatagagaagaacgcgcgagtatcacattcaattcaagaacaattcttg	1948
Dd	84548	ATTATATTTATTTAATTAAATTTATTAATTAATTTATTTATTTATTTAATTTAAATTAAT	84489
Qy	1949	aaaatgltataattcttcaacatactlaaanaatagtcctlaatgfatctccat	2008
Dd	84488	ATTATATTTATTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84429
Qy	2009	gtctctaaatatctttttttatatttagtctataatcatatgaaccaataacagt	2068
Dd	84428	TTTAATTTATTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84370
Qy	2069	gtgaatctcaaatcatccatcataatcttcttgaactcaacaattcaatcatlagca	2128
Dd	84369	TAAATTTAATTAATTTTATTTATTTAATTAATTAATTAATTAATTAATTTTCTATTTAT	84310
Qy	2129	ataacaatgcatagagaagttcccaaaaaaatttgttaacgaaatcccaattcttc	2188
Dd	84309	ATTTATTTAATTAATTTCAATTTTATTTAATTAATTAATTAATTAATTAATTAATTAAT	84250
Qy	2189	tcttcatgaaacaagaataacagatacagaacatacttctgtgtgbaatggagaagtaaa	2248
Dd	84249	ATTTAAATTTATTTATTTATTTATTTATTTATTTAATTAATTAATTAATTAATTAAT	84195
Qy	2249	tatcatcaagcaaatcttcaaaaaattcatate	2280
Dd	84194	TATATATATATTTAATTTAATTTAATTAATTAATTAATTAATTAATTAATTAATTAAT	84163
<b>RESULT 8</b>			
LOCUS	AE001398		
DEFINITION	Plasmidium falciparum chromosome 2, section 35 of 73 of the complete sequence.		
ACCESSION	AE001398		AE001362
VERSION	AE001398.1		GI:3845197
KEYWORDS			
SOURCE			
ORGANISM	Malaria parasite P. falciparum.		
REFERENCE	Plasmidium falciparum		
AUTHORS	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.		
	1 (bases 1 to 14867)		
	Gardner,M.J., Pateltein,H., Garucci,D.J., Cummings,L.M., Aravind,L., Koonin,E.V., Shallom,S., Mason,T., Yu,K., Fujii,C., Pederson,J., Shen,K., Jing,J., Aston,C., Lai,Z., Schwartz,D.C., Pertea,M., Salzberg,S., Zhou,L., Sutton,G.G., Clayton,R., White,O., Smith,H.O., Fraser,C.M., Hoffman,S.L., et.al.		
TITLE	Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum		
JOURNAL	Science 282 (5391), 1126-1132 (1998)		
MEDLINE	99021743		
REMARK	Erratum:[published erratum appears in Science 1998 Dec 4;282(5395):1827]		
REFERENCE	2 (bases 1 to 14867)		
AUTHORS	Gardner,M.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-NOV-1998) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20814, USA		
FEATURES			
source	Location/Qualifiers		
	I...14867		



[illegible]

D	b	7392	ATAATATATTTAAATATTAACCAAAATTAATATTTAAATTAATTAATTAATATATATTA	7451
O	y	406	tttggtctcgagtgaggttattcttgacaataatcacatgcacatctcttgattacta	465
D	b	7452	TATTAATATATTTATTTATTTATTTAAATTAATTAATATATTAATTAATTAATTA	7511
O	y	466	tgaagatttctc-----taccatthaatttcgcattcatatcctctgattataatt	520
D	b	7512	CAGCTTAATTTTAATATTAATTAATTAATTAATTTAAATTAATTAATTAATTAATTT	7571
O	y	521	aatacagatgctgataccgcttaccgaatcccaacatcagatcagatctctgctc	580
D	b	7572	TAAATTAATAAACAATAATATATGTGTTATTAATTAATTAATTAATTAATTAATTTCA	7631
O	y	581	aatccagcaaatlttatacaagaagatctgagaaaaaacccgaataaagaagggaa	640
D	b	7632	TATTAATATTTACTTAATAAATAAATTAATATAATTAATTTCAATTAATACATTAATTT	7691
O	y	641	gtatgacccatcggaagtatgtaataatcataagaagataagagatgacacccaag	700
D	b	7692	AATTAATATTAATAAATAATATTTTAA-ATTAATAATTAATTTTAAATTAATTAATTA	7750
O	y	701	gtctggaataatgctcccgacgctcttcacaaatcacatccgacccattgatt	760
D	b	7751	TATTTTATTAATTAATTAATTTAAATTAATTAATTAATTAATTTTAAATTAATAAT	7810
O	y	761	ttctcgataltcgltlaaaattcgcatacaagatctgtagaaaaatttatgttagct	820
D	b	7811	ATAAATATATTAATTAATTAATTTTAATTAATTAATTAATTAATTAATTAATAACAT	7870
O	y	821	gatctcaatal----tatgtccaggtatttgcaaatctctglttlaagaatcttgc	877
D	b	7871	TATATTTAAATTAATTAATAAGTAAATTAATAAATAAATTTCTTTTAAATTTATTTTAT	7930
O	y	878	cttctcttggctcglttctcttaactatatactacgcgatatagatacaatgat	937
D	b	7931	AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTAAATGTT	7990
O	y	938	atat-cacaaacacatgctcggacatcttgaataaacttctcacaacattacgg	996
D	b	7991	AAATTAATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATAATTAATTAAT	8050
O	y	997	aacactggaactcgacccttaaaatacgaatttaccagcgtaactagtgagattactagat	1056
D	b	8051	ATAATATATTTACATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTTATTT	8110
O	y	1057	aaagcataaaggaccgcttcaagctatttatacaaaagctacaaactgataagcttga	1116
D	b	8111	AAATGAATTAATAAATAATTAATTAATTAATTAATTTAA--TTTATTAATTAATAATTAATTA	8167
O	y	1117	atcccttgaataaacttgggaattaccggtgtgatgtaaatatagaattagtggtgaaca	1176
D	b	8168	TATATTAATTAATTAATTAATTAATTTATTTTAAATACTTAATTAATTAATTAATTAATTT	8227
O	y	1177	aatabgtthaatcaatctagtggtcaacatacacaatctcccttaccgaaaaaacaactta	1236
D	b	8228	ATAAAAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	8287
O	y	1237	agagaagtatacatcatataggtaagctatatacccttcaagtgatctactagag	1296
D	b	8288	ATTAATAATTAATAAATAATTAATTAATTTTAAATTAATTAATTAATTAATTAATTAAG	8347
O	y	1297	actaaagaatgctatgctgctgcataaatagaatctcacacgcyggaatcaattacg	1356
D	b	8348	TCTATTTTAT--TTATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT	8405
O	y	1357	gaccgatgctacatcacatcgaataatcatctctggttggttcaacataaacaacaaa	1416
D	b	8406	TTGTTTCTTTCTATTAATTAATTAATTAATTTTATTTATTAATTAATTAATTAATTAATTA	8465
O	y	1417	caagaaaaaagaacagatttctctggaatcccatcacaagatcatataaagtcatagct	1476
D	b	8466	TTAATAAATAATTAATAATTAATTAATTAATTTTAAATTAATTAATTAATTAATAAATAATTA	8525

OY	1477	cttggggttaacagtttcggaagctcctcaagaagcggttcaacccctcgcaacttcaaatcgt	1536
DB	8526	tttccaatatataatataatataatgcattttttttaaagtttaatatattatttat - ttaattttt	8584
OY	1537	cttcttcaatgcacttcaacatatttctgttgaattggaattcaatgaagcgaaact	1596
DB	8585	atataaataataatattgcttttaattcatttaattataattataatttttaataataata	8644
OY	1597	gtaacattacaatcttataatagtaacagtaatgtaatgattatccaaataacttcttg	1656
DB	8645	attatgatataatattttatattatttatataatttaattagctatataataaatttaacatat	8704
OY	1657	algttaaaacttaacttgccttctccctcaagatataaataltaatcatcgaggtataaaa	1716
DB	8705	acattatataatttaacatatattttattatttttaatttaatttttaatttttaattttt	8764
OY	1717	agtttgcattatttcgcagatcgaaaggaatgaaccacaaagacttaatttttgaa	1776
DB	8765	ttaattatt	8824
OY	1777	atgtaaaccttcttcaatagatataatccagatgttctgttccaaatagcaagcct	1836
DB	8825	ttatt	8883
OY	1837	ctcaacacgtgatgcataatttttctgcaaatataatgsgaatccaatgcatact	1896
DB	8884	attttatt	8943
OY	1897	caatagaagaacacgcgtgaglatcaacttcaattcaagaacaaaacttttgaaaaagt	1956
DB	8944	ttatt	9003
OY	1957	tatatcttcaacacatcttcaaatatagagcccaatgattctccatgctcttaa	2016
DB	9004	tatt	9063
OY	2017	aatacttttttttatacttgatcataaatacatlalgaacaaataatgctgg ----t	2070
DB	9064	att	9123
OY	2071	gaattcaaatatcccatataatcttggaaatccacaatatataatattgtaact	2130
DB	9124	tatt	9183
OY	2131	aacaaatgcatagaagttccaaaaaaaactttgtaacgaacattccaaatttttct	2190
DB	9184	tatt	9243
OY	2191	tttctggaacagaataaacaagatagaacaaacttttctgtggaatggaagtgtaata	2250
DB	9244	tt	9303
OY	2251	tacattgaagcaaatcttcaaaaaatataataagctataacgcgtccaaagtatcatca	2310
DB	9304	tatt	9363
OY	2311	gttagtgtaattataatgatgatggtgcgtatccgaattgggaacaaatgaaacgggaat	2370
DB	9364	tt	9421
OY	2371	taaaatacttaactttaaaataaataaaatctgagtaaatggtcttctgcatactggg	2430
DB	9422	acacatt	9481
OY	2431	ggcaaaaaaagaacatgccaagaactca 2459	
DB	9482	acttaaaaaatttaataatataatataatgata 9510	

RESULT	9			
AC025417				
LOCUS	AC025417	106142 bp	DNA	PLN
				09-JAN-2001

DEFINITION	Genomic sequence for Arabidopsis thaliana BAC T12C24 from chromosome I, complete sequence.
ACCESSION	AC025417
VERSION	AC025417.4
KEYWORDS	HTG.
SOURCE	GI:9438236
ORGANISM	thale cress. Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 106142) Chao,Q., Brooks,S., Buehler,E., Johnson-Hopson,C., Khan,S., Kim,C., Shin,M., Altaifi,H., Bel,Q., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharisky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vaysberg,M., Yu,G., Federspiel,N.A., Thaveri,A., Tortum,M., Vaysberg,M., Yu,G., Federspiel,N.A., and Ecker,J.R.
TITLE	Genomic sequence for Arabidopsis thaliana BAC T12C24 from chromosome I
JOURNAL	Unpublished
AUTHORS	2 (bases 1 to 106142) Ecker,J.R.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAR-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE	3 (bases 1 to 106142) Ecker,J.R.
AUTHORS	Direct Submission
TITLE	Submitted (28-JULY-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
JOURNAL	4 (bases 1 to 106142) Chauk,R., Shin,M., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altaifi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharisky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., and Ecker,J.R.
TITLE	Direct Submission
JOURNAL	Submitted (27-JULY-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
REFERENCE	5 (bases 1 to 106142) Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altaifi,H., Bel,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howng,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharisky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Tortum,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., and Ecker,J.R.
AUTHORS	Direct Submission
TITLE	Submitted (09-JAN-2001) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
JOURNAL	On Jul 25, 2000 this sequence version replaced gi:7940297.
COMMENT	Location/Qualifiers
FEATURES	1..106142 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" /clone="T12C24" 121..786 /note="similar to CLV1 receptor kinase like protein emb1CA18252.1" /codon_start=1 /evidence=not experimental /product="T12C24.1" /protein_id="AAP8073.1" /db_xref="GI:9502366" /translation="MGLILSEFPNGSLVDNLRLRIPTGSSSYGNTDLNMRHFOIA
CDS	







YELKNHKEINRLKLYKKRAPHYIYVNDQFIINNYIIHKKIKHIIHILHKKCKRKE  
IDIVKREKDIYTCMIIINNINNIFFSIYHVDHNECVYRIFPKANKILLEYINLKNKEIY  
YLNKIKRYLYLNNNISNDLYPNVISYNRKIYMONKRNKRNFSEHIEYSIKNDIHLLE  
LYTORLONCKDEIYSYIKRYNFENKPNLYLVELYHIVYSKKRKKKFFESISSP  
YSDIVWNSFETSEYFFFSYLLEFIYHDMVASYIFPFTLYISLPPYYYSLNMI  
MVYCGPKT"  
join(7669, .7717,7800, .7829,7912, .7940,8064, .8106,  
8302, .8368,8568, .8641,8713, .8812,8924, .9006,9122, .9258,  
9369, .9505,9613, .9838)  
/gene="PFC0581w"  
join(7669, .7717,7800, .7829,7912, .7940,8064, .8106,  
8302, .8368,8568, .8641,8713, .8812,8924, .9006,9122, .9258,  
9369, .9505,9613, .9838)  
/gene="PFC0581w"  
/note="PFC0581w, hypothetical protein, len: 324 aa,  
revised: new gene prediction, splicing prediction very  
tentative"  
/codon\_start=1  
/product="hypothetical protein, PFC0581w"  
/protein\_id="CAB90285.1"  
/db\_xref="GI:7711066"  
/translation="MKKKKHRTSIFVAVDKKIVYLCSDYRIIRINYMIGHIFRL  
ONIFHTYFANEQIASFEFSHKYRVTKEDIIDIGICEKNWITDYLSIESIKDNDFS  
NOIKHTYFANEQIASFEFSHKYRVTKEDIIDIGICEKNWITDYLSIESIKDNDFS  
LTKKDEIYSKKNKMDQLLTTSYNNKIDVWASFOAQSPEYIFLNIKSHRSPGSA  
LYVEKKEKIIYKKNKMDQLLTTSYNNKIDVWASFOAQSPEYIFLNIKSHRSPGSA  
LYVEKKEKIIYKKNKMDQLLTTSYNNKIDVWASFOAQSPEYIFLNIKSHRSPGSA  
ONIKREKNNKMDQLDIDDEKRYFDEILRAKKKSEYDDBDEL"  
7790, .7799  
/note="potential splice acceptor sequence for exon 2 of  
PFC0581w"  
7830, .7835  
/note="potential splice donor sequence for exon 2 of  
PFC0581w"  
7901, .7911  
/note="potential splice acceptor sequence for exon 3 of  
PFC0581w"  
7941, .7947  
/note="potential splice donor sequence for exon 3 of  
PFC0581w"  
8055, .8063  
/note="potential splice acceptor sequence for exon 4 of  
PFC0581w"  
8107, .8112  
/note="potential splice donor sequence for exon 4 of  
PFC0581w"  
8293, .8301  
/note="potential splice acceptor sequence for exon 5 of  
PFC0581w"  
8369, .8374  
/note="potential splice donor sequence for exon 5 of  
PFC0581w"  
8559, .8567  
/note="potential splice acceptor sequence for exon 6 of  
PFC0581w"  
8642, .8647  
/note="potential splice donor sequence for exon 6 of  
PFC0581w"  
8704, .8712  
/note="potential splice acceptor sequence for exon 7 of  
PFC0581w"  
8813, .8818  
/note="potential splice donor sequence for exon 7 of  
PFC0581w"  
8913, .8923  
/note="potential splice acceptor sequence for exon 8 of  
PFC0581w"  
9007, .9012  
/note="potential splice donor sequence for exon 8 of  
PFC0581w"  
9112, .9121  
/note="potential splice acceptor sequence for exon 9 of  
PFC0581w"  
9259, .9264

[illegible]

[illegible][illegible]

RESULT	11
PfPRAL3P5/c	
LOCUS	PfMAL3P5      86827 bp      DNA      INV      04-MAY-2000
DEFINITION	Plasmodium falciparum MAL3P5, complete sequence.
ACCESSION	AF034556 AL008971 AF008972 AL008978 AL010141 AL010153 AL010162
VERSION	AL010206 AL010210 AL133179 AL034556.3 GI:7711064
KEYWORDS	HG; centromere; CTRF protein; initiation factor E4; Serine/threonine protein phosphatase. <i>malaria parasite P. falciparum.</i>
SOURCE	Plasmodium falciparum
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE	1 (bases 1 to 86827). Bowman,S., Lawson,D., Basham,D., Brown,D., Chillingworth,T., Churcher,C.M., Craig,A., Davies,R.M., Devlin,K., Felwell,T., Genillies,S., Gwilliam,R., Hamblin,N., Harris,D., Hayford,S., Hornaby,T., Horrocks,P., Jagsal,B., Jessel,B., Kyes,S., McLean,J., Moule,S., Mungall,K., Murphy,L., Oliver,K., Quail,M.A., Rajandream,M.-A., Rutter,S., Skellton,J., Squares,R., Squares,S., Sultson,J.E., Whitehead,S., Woodward,J.R., Newbold,C. and Barrell,B.G.
TITLE	The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum
JOURNAL	Nature 400 (6744), 532-538 (1999)
MEDLINE	99376085
REFERENCE	2 (bases 1 to 86827) Bowman,S., Skellton,J., Churcher,C., Lawson,D., Quail,M. and Barrell,B.

JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 86827)  
AUTHORS Lawson, D., Bowman, S. and Barrell, B.  
TITLE Direct Submission  
JOURNAL Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium,  
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge  
CB10 1SA, UK  
COMMENT On May 5, 2000 this sequence version replaced gi:4493931.  
For more information about this sequence or the Malaria Project,  
see [http://www.sanger.ac.uk/Projects/P\\_falciparum](http://www.sanger.ac.uk/Projects/P_falciparum).  
Location/Qualifiers  
1. 86827  
/organism="Plasmodium falciparum"  
/strain="3D7"  
/db\_xref="taxon:5833"  
/chromosome="3"  
/clone="MAL3P5"  
join(322..603,826..1023,1197..1301,1458..2942)  
/gene="PFC0575w, MAL3P5.1"  
join(322..603,826..1023,1197..1301,1458..2942)  
/gene="PFC0575w, MAL3P5.1"  
/note="PFC0575w (MAL3P5.1), Hypothetical protein, len: 689  
aa, possible signal sequence, revised: added new exon 2"  
/codon\_start=1  
/product="Hypothetical protein, PFC0575w"  
/protein\_id="CAB38969.2"  
/db\_xref="GI:7711065"  
/db\_xref="SPTREMBL:O97258"  
/translation="MYLKNVYIYISCFILFDLCFSFHLMKRYKNNMNMKSVTFL  
RSPOYIRKRKRSRIRKNSVFKKOKKPLFENLKQCSFLGWRNOYDQYIDVIS  
NINNLRIKQVTKHKKSNFTKENIKQILHCYFSKIDFKIINLSYIIRHQSNT  
VHSILNOISEKVEKDAENYLAHLFLKDENTLFSMHIMDFEFSKQVIECID  
IKSKKKRKLSTIYNLFICTLIYFCMLLTKIKYISLIIFFEFPCFLCNILIER  
IYECVGDILRKIRERNLYCEKKIKFHKADAKKKEITMKDDIYFNHYDELLIC  
PYKLIERNKNNKNIIRSNYNINDISIDOMTNNNTIDVNNISIDEKIKDEFP  
DDENLKLDYEQFLQENDNIIKYIEEDPLYNINNSNINNNININIKKHKIK  
DYNDDDDYEEKEDLYIQKNDIDYIKNTIGNNKLEFEKQFIQADIEFQNS  
NVNLDGKRVKSNDEKSTEHKNTINKGDTGLQOMENNFIRKNDINISD  
ISNNEINIKLKNOSDEDINLSDIYERLRKRYLWYIYOKIEYKFOYDITND  
YPIIKNKTYLDLNTGYKIVMSPDVNSLFEKTKIDISINERKDNQNMOKSKRY  
TIMPNWY"  
604..609  
/note="predicted splice donor sequence for exon 1 of  
PFC0575w (revised)"  
816..825  
/note="predicted splice acceptor sequence for exon 2  
(revised of PFC0575w)"  
1024..1029  
/note="predicted splice donor sequence for exon 2 of  
PFC0575w (revised)"  
complement(3354..6644)  
/gene="PFC0580c, MAL3P5.2"  
complement(3354..6644)  
/gene="PFC0580c, MAL3P5.2"  
/note="PFC0580c (MAL3P5.2), Hypothetical protein, len:  
1097 aa, possible signal sequence, predicted using  
hexExon"  
/codon\_start=1  
/product="Hypothetical protein, PFC0580c"  
/protein\_id="CAB38968.1"  
/db\_xref="GI:4493932"  
/db\_xref="SPTREMBL:O97257"  
/translation="MGILKRNVPYLVSPSLFAYFLKHKNNENYELLINNDIE  
KIKIRIHNKCSYIPLFLNIYDSYIKKKILRLNLYKFRKRRDKDEYIITMVK  
KRRRAIKYNEISDEQNLNFKFYIEVLEYSLKGLISPHLSYILKNISEHCNIVK  
SLYYNKLNNKHLINERKLYEFOKINNEHTQOAPTHHTNNNNNNKRPIDIIHSC  
KNTNISYSTYNNNEKENINIDKYNIHNFYTERKISYKDCOHITLMNYLNDQY  
DNICRIQNTNININININILKYCYKNNMETIILNYNHIEEMKKKINOKNTNYS  
LFYIYSFPEFKRENHIYDLFEQMMNHLKKNDEKPYNSNNTNNNTIKYISDW  
FYHINSSNRCSFKNLKQOTDNTNKHITMGKRYPMNSDHEKNNNTGCGNINIK  
DQKDIILAKIYFLKGNKLDIQLINELVMIYMLLFECSLISIKKNIHLLEKKE  
FDKDNKIYINSADYNNNLRNLIKRSKNEERINISFASFPILSKNIYFDEIG  
RSRDNITYNVYDEKTNKTTNNNNNDNDNICSNNDHICSNNDHICSNNDHICSN

NNNNICSNNNNNICSNNNNNICSNNNNNICSNKMLDFFCODKNFNDYNTKKKKRRI  
YELAKITNFIPLFKGKKEKHONEDNTIMLYIKKEYPWIFLYKNIIRKKDPSFIEH  
NNNIVNDGIDTKNNNIIPEKKRYNLESSISFYIKDIEYVYKRLAYNLIKKRGR  
YFLKMNHIRKRLKMKRAHYIYNDDQITINNYIITIKKNIHKHILHKQCKDE  
IDIVKFDLIYCCNINININISFYIHKDHNHCYRIFKAVNKLILEYENLNKENY  
YKNNIKKRYLTLNNNISNDLYPYNISYKRYNOKKYNRNFNSHIFSLNDHLFL  
LYTORIQCDDISFYIKRYNFKNEKFLYLYELHYIYSEKKKKKRPFSSSD  
YSYDTONSFTFSYFFFSYSLYFLFLEPHDWMASYIFFFTLYSGLPYYSLYNNI  
MYVCGPRT"  
join(7669..7717,7800..7829,7912..7940,8064..8106,  
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,  
9369..9505,9613..9838)  
/gene="PFC0581w"  
join(7669..7717,7800..7829,7912..7940,8064..8106,  
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,  
9369..9505,9613..9838)  
/gene="PFC0581w"  
/note="PFC0581w, hypothetical protein, len: 324 aa,  
revised: new gene prediction, splicing prediction very  
tentative"  
/codon\_start=1  
/product="hypothetical protein, PFC0581w"  
/protein\_id="CAB90285.1"  
/db\_xref="GI:7711066"  
/translation="MKKKKHYSIKFVNDKRYVLCSDYIRIINWIGLHIFRL  
OKIHFYVANDQIASFFESHKRYRVRKEDIQIECKWNIDYLSIESIKODNS  
NDIKTYTAKNNKNDOLLTSYSKRKIDTYNASQWQOSEYIFLNKRFHSRNSPRA  
LKYNDEIVSKKNFSALSNDNSYTKYIDYLDLNIISERTYNASGVKVV  
TLKREKRIINRLLSKERYPNQVMDMEKRIHITPTVINLFFLSISHRHVS  
QNFLEKEKNNSDKLQDDIDDEEKYFDEILREAKKSEEDKDEEL"  
7790..7799  
/note="potential splice acceptor sequence for exon 2 of  
PFC0581w"  
7830..7835  
/note="potential splice donor sequence for exon 2 of  
PFC0581w"  
7901..7911  
/note="potential splice acceptor sequence for exon 3 of  
PFC0581w"  
7941..7947  
/note="potential splice donor sequence for exon 3 of  
PFC0581w"  
8055..8063  
/note="potential splice acceptor sequence for exon 4 of  
PFC0581w"  
8107..8112  
/note="potential splice donor sequence for exon 4 of  
PFC0581w"  
8293..8301  
/note="potential splice acceptor sequence for exon 5 of  
PFC0581w"  
8369..8374  
/note="potential splice donor sequence for exon 5 of  
PFC0581w"  
8559..8567  
/note="potential splice acceptor sequence for exon 6 of  
PFC0581w"  
8647..8647  
/note="potential splice donor sequence for exon 6 of  
PFC0581w"  
8704..8712  
/note="potential splice acceptor sequence for exon 7 of  
PFC0581w"  
8813..8818  
/note="potential splice donor sequence for exon 7 of  
PFC0581w"  
8913..8923  
/note="potential splice acceptor sequence for exon 8 of  
PFC0581w"  
9007..9012  
/note="potential splice donor sequence for exon 8 of  
PFC0581w"  
9112..9121  
misc\_feature











REFERENCE  
AUTHORS  
TITLE  
2 (bases 5269 to 5695)  
Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.  
Transfer RNA genes in *Drosophila* mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial tRNA genes  
Nucleic Acids Res. 11 (8), 2411-2425 (1983)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
3 (bases 404 to 5272)  
de Bruijn,M.H.  
*Drosophila* melanogaster mitochondrial DNA, a novel organization and genetic code  
Nature 304 (5923), 234-241 (1983)

JOURNAL  
MEDLINE  
AUTHORS  
TITLE  
4 (bases 804 to 1778)  
Satta,Y., Ishiwa,H. and Chigusa,S.I.  
Analysis of nucleotide substitutions of mitochondrial DNAs in *Drosophila* melanogaster and its sibling species  
Mol. Biol. Evol. 4 (6), 638-650 (1987)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
5 (bases 5268 to 13619)  
Garesse,R.  
*Drosophila* melanogaster mitochondrial DNA: gene organization and evolutionary considerations  
Genetics 118 (4), 649-663 (1988)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
6 (bases 441 to 2967)  
Satta,Y. and Takahata,N.  
Evolution of *Drosophila* mitochondrial DNA and the history of the melanogaster subgroup  
Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
7 (bases 14215 to 14512)  
Ballard,J.W., Olsen,G.J., Falth,D.P., Odgers,W.A., Rowell,D.M. and Atkinson,P.W.  
Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods  
Science 258 (5086), 1345-1348 (1992)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
8 (bases 14917 to 19517)  
Lewis,D.L., Farr,C.L., Farguham,A.L. and Kaguni,L.S.  
Sequence, organization, and evolution of the A+T region of *Drosophila* melanogaster mitochondrial DNA  
Mol. Biol. Evol. 11 (3), 523-538 (1994)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
9 (bases 1 to 408; 13319 to 19517)  
Lewis,D.L., Farr,C.L. and Kaguni,L.S.  
*Drosophila* melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons  
Insect Mol. Biol. 4 (4), 263-278 (1995)

JOURNAL  
MEDLINE  
REFERENCE  
AUTHORS  
TITLE  
10 (bases 1 to 19517)  
Lewis,D.L., Farr,C.L. and Kaguni,L.S.  
Direct Submission  
Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA  
Location/Qualifiers

FEATURES  
source  
1. 19517  
/organism="Drosophila melanogaster"  
/organelle="mitochondrion"  
/db\_xref="taxon:7227"  
/note="derived from new and previously submitted sequences: sequence is a composite containing sequences obtained from different *Drosophila* melanogaster strains"

1. 65  
/product="tRNA-Ile"  
complement(97. .165)  
/product="tRNA-Gln"  
171. .239  
/product="tRNA-Phe"  
240. .1265  
/codon\_start=1  
/transl\_table=5  
/product="NADH dehydrogenase subunit 2"

CDS

/protein\_id="AAC47811.1"  
/db\_xref="GI:1166530"  
/translation="MPNNSSKIFLETTIMICTLTYSNSGLAMGLEINLSFPL  
LSDNNMLSTEASLKYFLVYLASTVLLFSSITLMLKNNMNEINSEFTSITMSALL  
LKSGAAPFHEWFMMEGLTMMNALMTWKIAPLMLISLNTKYLILLSVLSVIT  
GAIGLNGTSLRKLMARFSSINHLGMLSSIMLSIMLILEFFSFKSFLYTFEWF  
KLPHNOLSEWFSVNSKILKFTLPNPLSLGLPELPKPLVIOOLTCONOYEMLT  
IMMSLTITLFEFLRYICYSAPFMNYPENNIMKNNMSININMIMTFPSIRPLFLI  
SLFFMF"  
1264. .1329  
/product="tRNA-Trp"  
complement(1332. .1383)  
/product="tRNA-Cys"  
complement(1403. .1468)  
/product="tRNA-Tyr"  
join(1470. .1472,1474. .3009)  
/note="mechanism underlying reading frame shift after first codon uncertain"  
/codon\_start=1  
/transl\_table=5  
/product="cytochrome c oxidase subunit I"  
/protein\_id="AAC47812.2"  
/db\_xref="GI:7412849"  
/translation="MSRWLESTNKKDITLYEFGAMAGVGTSLILTRAEIGHPC  
ALIGDDIYVIVTAHAFAFMIEFMVPMIGFGNMIVPLMGAPDMAFPMNMSFW  
LLPALSILLYSSWVAGAGGTVPYPLSAGINAGSADIATFSHLNGISILGA  
VNFITTYNNRSTGISLDRMPLFVWSVITALLLSLSPYAGATIMLIDRLNLTSE  
FDPAGGDPILYQHLWFEGHPEVYILILPEFGMISHISOESGKRTFSLGMIYAM  
LAIDLGEIVAHMFTVGMVDVRAVTSATMIJAVPTGKIFSMATLHGDLSTYS  
PAIIMAGFEVFLFTVGILGVLANSDITLHPDYVAHFHYLSMGAVFAIMAF  
IHWPFLFTGILNNKMKSHETIMFIGNVLFPGHGLAGMPRYSDDPATYTN  
IVSITGISTISLGLTFEFLTWESLVQRQVITYIQLNSTIEWQNTPRAHSSELP  
LITN"  
3012. .3077  
/product="tRNA-Leu"  
3083. .3767  
/note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"  
/codon\_start=1  
/transl\_except="(pos:3767,aa:TERM)  
/transl\_table=5  
/product="cytochrome c oxidase subunit II"  
/protein\_id="AAC47813.1"  
/db\_xref="GI:1166532"  
/translation="MSTWANGLDSDASPLMEQILFFHDHALLILWATVYGLMER  
LEFNNYVNRFLHGLIEMITLIPAILLEFLALPSRLYLLEIDNEPSYTLKSIGH  
QWVSYSESDENLEPDSYMTPLNELMTDGRLLDVNRVVLPMNSQIRILVTAADVI  
HSMVPAIGVAVDGTGRLNQTNFNINPGILFYGCSCIGANHSFMPVIESVAVY  
FIRKISSNNS"  
3768. .3838  
/product="tRNA-Lys"  
3840. .3906  
/product="tRNA-Asp"  
3907. .4068  
/codon\_start=1  
/transl\_table=5  
/product="tRNA-Phe"  
/protein\_id="AAC47814.1"  
/db\_xref="GI:1166533"  
/translation="MPQMAPISWLLFLIFSTPLRCSINYSYMPNPSPKSNELKNI  
NLNSNMNKM"  
4062. .4736  
/codon\_start=1  
/transl\_table=5  
/product="tRNA-Pase 6"  
/protein\_id="AAC47815.1"  
/db\_xref="GI:1166534"  
/translation="MMNTNLSVDPDLAIFNPLSLMWLSTPLGLAMPSPYIWMPSRYNI  
MMNSIILFLHKEFTLILGSPSCHNSSTFISLSPLILFNNMGFLPIFTSTSLITLT  
LSLALPLMLCPMLGTNINHTOHFAHLVPGSTPAILMPFVCIETISNITPGLTAR  
LTANMIGHLTLTLTGNTGSSMSYMLTFLMAOIALLVLESAAVAMIOISYVFAVLSPL  
YSSFVN"  
4736. .5524  
CDS



1607 aatattatataagatactagatgatttcccaataacactcttgatggtttaaac 1666  
18737 AATTAAAAAGATGATATTAATTTAAAAATTTATATTCATATTTATTTATTTAA 18796  
1667 ttaactctgttctcctcagataataatatacatcagaggttaaataagtttgtct 1726  
18797 TTTAATTTATATTAATTAATTAATTAATTAATTTATTTATTTATTTAAATTTATAT 18856  
1727 tatttcgcgagtcagagataaactaatgactttaattttttgaaatgttaacct 1786  
18857 ATTATTTGAATTTATTAATTAATTAATTAATTAATTTATTTATTTAAATTA 18915  
1787 cttaactcagataaactaacgcatgttttggttcgataagagcctcacacgct 1846  
18916 TTTAATTAATTAATTTTAAAAATTTCTTAATGTTATTTTAAATTAATTTATAT 18975  
1847 gatagtcaattttctcgaataataatgaatgaatcaatgctactacataagaaga 1906  
18976 AATTAATCAATTTTATTTAAAAATTAACAAAAATTTTAAATTAATTAATTTAAATG 19035  
1907 aacgcgtgagatcattcaatttaagaacaaattttgaaataatgtaattctc 1966  
19036 AAATTAATTTATTTATTTTTCATTTTAAAAAAATTTTAAAAAAATTTATTTT 19095  
1967 aacaatattataaataatgagtcataatgattctcctatgctcttaaaatatttt 2026  
19096 TTTTAAAAAAAACATATTAATTAATTAATTAATTAATTAATTTATTTATTTATTAAT 19155  
2027 ttatatttgatataaataacatataatgaacataatggttgatgaatcaatactcc 2086  
19156 TTAATATATTTATTAATTAATTAATTAATTAATTAATTAATTTAAATTTAAATTTAG 19215  
2087 attaatatttttgtaaacctacaataatataatttgatgaatcaatgataagaag 2146  
19216 AATTAATTTATTAATTAATTTATTTTTCATTTTAAATTTATTTATTTATTAATTA 19275  
2147 ttcccaaaaaaatttgtaagaagaacctccaatttttttttttgtaagaagaaga 2206  
19276 TAAATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 19335  
2207 taacagataagaacataatttggttgtaagaagatgaatatacatataagaacatt 2266  
19336 TATATTAATTAATTAATTAATTAATTAATTAATTTATTTAAATTAATTTTTCATTAATA 19395  
2267 taataaatatataagcctatacgccgctcaaaagtgtatcagtgagtgtaataata 2326  
19396 AAAAAATTTCTTAATTAATTAATTTATTTATTTATTTATTTATTTATTTATTTAA 19455  
2337 atgcagtgtgcattcgaattggagacaacatgaacaggaatttaaatattacttta 2386  
19456 TTTTAAAAAAAATGATTAATTAATTAATTAATTAATTAATTAATTAATTTATTTA 19515

RESULT 14  
PFMALIP3/c  
LOCUS PFMALIP3 67970 bp DNA INV 15-DEC-1999  
DEFINITION Plasmodium falciparum MALLIP3, complete sequence.  
ACCESSION AL031746  
VERSION AL031746.9 GI:6594243  
KEYWORDS HTG.  
SOURCE malaria parasite P. falciparum.  
ORGANISM Plasmodium falciparum  
REFERENCE 1 (bases 1 to 67970)  
AUTHORS Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M. and Barrett, B.  
TITLE Direct Submission  
JOURNAL Submitted (24-SEP-1998) P. falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK  
COMMENT On Dec 16, 1999 this sequence version replaced gi:5763807.  
For more information about this sequence or the Malaria Project,

see [http://www.sanger.ac.uk/Projects/P\\_falciparum](http://www.sanger.ac.uk/Projects/P_falciparum). IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E. coli, yeast, vector, phage etc.

## FEATURES

Source  
1. 67970  
/organism="Plasmodium falciparum"  
/strain="3D7"  
/db\_xref="taxon:5833"  
/chromosome="1"  
complement(join(1748. 2598,2748. 2848,2990. 3276))  
/gene="MALLIP3.01"  
complement(join(1748. 2598,2748. 2848,2990. 3276))  
/note="MALLIP3.01, conserved hypothetical protein, len: 412 aa, similarity: UPF0006 family eg to YB1055C/YB10512/YB10511, YB5, YEAST (418 aa), fasta scores: opt: 316, E(): 1.1e-12, (33.28 identity in 271 aa overlap)"  
/codon\_start=1  
/product="conserved hypothetical protein, UPF0006 family"  
/protein\_id="CAB63556.1"  
/db\_xref="GI:6594244"  
/translation="MKLVFHYIKYINLFIYSIFLKSNSKITNDLRYSTVKKYV LQIKRSLMKNNHNRKMEDESSFDIGSLWLTQKMGVNSKHEHDLONVNRK NNNVDKIITTCYLAELIDKSLKICEYDPEKFLYLSAGVPHNCEYEDIDNKEBE IIAKREYEEPIKFAENQVENSKEGNGKRCDCDEKMMNNINELLENDTIPGKY NEKREYLENKNIKIKYPRIVCIGEGELDFEISKFIQKVFQKLVQVGM LPMFLHMGNESEPFKTIYDIYKFLPERKNGVTHSETTKEDIVHIVQNYKLYGVNG CSLSLEINAVKKIPIPLLLELIDAPMGCGVKTHASIEYIKDYEKATYLNKIKIN IICDDNTIFERENPEYNTIA  
complement(2599. 2610)  
/note="potential splice acceptor sequence"  
complement(2742. 2747)  
/note="potential splice donor sequence, atg/gttaa"  
complement(2849. 2861)  
/note="potential splice acceptor sequence"  
complement(2984. 2989)  
/note="potential splice donor sequence, aaa/gtaaaa"  
5005. 5496  
/gene="MALLIP3.02"  
5005. 5496  
/gene="MALLIP3.02, hypothetical protein, len: 163 aa, contains possible signal sequence"  
/codon\_start=1  
/product="hypothetical protein, MALLIP3.02"  
/protein\_id="CAB63557.1"  
/db\_xref="GI:6594245"  
/translation="MKLLNRFVVCPIIILFFFLNSVLCNNRRNINIFETENAK AMRKLSEINSIKLDNGDELKIKLNDKHKDSYKWSYSEFISNLEEKYSQDLEFK KKOINEANTKIIDROFYIINDIEINIAITRYLENPFELYIQSPKQSLIDIIOS LNN"  
8020. 10389  
/note="possible cen1, region of very high [A+T] content"  
14884. 20352  
/gene="MALLIP3.03"  
14884. 20352  
/gene="MALLIP3.03"  
/note="MALLIP3.03, putative ABC transporter, len: 1822 aa"  
/codon\_start=1  
/product="putative ABC transporter"  
/protein\_id="CAB63558.1"  
/db\_xref="GI:6594246"  
/translation="MTYKKENVGTSNGKNGKKKSSCONISFLNLSFPMIRPLINDLIK GDIOELPNICNPVPVYASKLEENLRDIEVDEPFSSEKSSNEHVIHRNSNDASE KKVNYIHNILWISLITFKRILIIISFIIITLFLVLTGAKFIDYIMRLIEGQKIP YISFLKDFKFSGLVVMIMFHLFFALHLYPHLFTINLKVSLMFLVKIKNSNN NHIQNDPAFYNTYRRFSQDEISDFLSIGFNASSSSGICNNKNNKNNKNNKNNFEVN DYIINFKSTKMKKDSLNERSLPNNYIYIMSDVPSVTFYFSCINLENVVKLF



OY	556	caacatgattgatgccttgctgtcctaaccagccaattattaacaaggatltgagaaa	615
Dd	9855	ATTAATAATTATTTATTTAATAATAAAAATAATATATATATTTATATATTTAAATTAAT	9836
OY	616	aaccgaataataagaagaagtagtgcaccaatggagtatgtaatatatacaa	675
Dd	9835	ATATTTTATTTTTAAAAATATAATTAATTAATAAATGCAATATAATATACGTATATA	9776
OY	676	gagataagagatgacaaccaaaagcttgtgataataygtccctcgcaagcttcctca	735
Dd	9775	TACAATTAATAATAATAATTTTAAAGCTAATTTAAAT - ATATATATATTTTAAATA	9717
OY	736	caatcatatcacgccactttggaattcttcgataatctgttaaattcgataagat	795
Dd	9716	TTATTTAATATATATTTATATAAAAAT-----ATATATATATTTAAAAATTTTATA	9662
OY	796	gtgaaaaaatatttatltgtltagctgcatacatatgctccagatattgcataac	855
Dd	9661	TTTATATATATTCATATATAAATAATATATATATATATATATATATATATATTA	9602
OY	856	tctcgcttaagcatatttgctcttcctttgcttgcttcgttccttaactatatac	915
Dd	9601	ATTTATTTAATTAATTAATAATATATTTATTAATTTATATATATTTTAAAAATATA	9542
OY	916	gcggatalatgatatacaatgalataatacaaaaacattgctcggaacctiltgaataa	975
Dd	9541	AAATAAATTAATATATATTTTAAATATTTATAAGAATATATTAATAATAATTAAT	9482
OY	976	cctttctcaaaccttacgggacacgcggactgcgaaccttaaatcgaattttagcgc	1035
Dd	9481	AAATGTGAATTTATTTATTTAAAAATATATATTAATAA- AATATATATTAATAATAATGAA	9423
OY	1036	actagttgagatlaactagcaataagataagaaccgcgticaagctattatacaagt	1095
Dd	9422	ATATTTATATATATAAATAATATATATATATATATTTGTTTTATTTAATAATATATA	9363
OY	1096	acaaactgaaatagcttgaaccccttgaanaatttlynaatccggitgtatglaa	1155
Dd	9362	ATTATATATATATATATATTTATTTATTTAAAGTGACTCAAATATAAATAATTAATAT	9303
OY	1156	atataagctttagtgtaacaaatlatgatacoaactttagtgcacacatacaaatc	1215
Dd	9302	ATATATTTATATATATTTAAATAATAATTAATTTATATATTTAATTAATAATTTATA	9243
OY	1216	cttaacagaaaaacaaacttaagaagatbaacataatccataatayggulatgctaect	1275
Dd	9242	ATTTATTTATTTATTTATTTATTTATATTTATTTATTTATTTATTTATTTATTTATTT	9183
OY	1276	ttaacgtatgctatactagaagactaa-agaaatgltatgtatgctgcogataatgaatic	1334
Dd	9182	ATATATTTAAATTAATTTATTAATTTATTTAAATTTATTTATTTATTTATTTATTTAATA	9123
OY	1335	aacagcggtgaaatcatlacgggacgclatgtaacgaacgaacaaatcttcgt	1394
Dd	9122	TTTATATTAATAATAATTTGTTTTATTTATTTAATAATTTATTTATTAATAATTTTATAT	9063
OY	1395	tggtcaacaataaaaacaaaacaganaa----aaegaaaacgatlltctlgatl	1448
Dd	9062	TTTAAATTTATTTATTAATAATTAATAAGCATATATTTAAACATATAATTAATATATAT	9003
OY	1449	ccattcaatgtctcaaaatgcatagatcctttgggttacaagtltcogaagctcctacaag	1508
Dd	9002	TTTATTTTATTTTATTAATTTGGTTTAATATATATTTATTTTATTTTAAATTAATTTTATAT	8943
OY	1509	cgtttaaccacitgcaactiaaattgccttccttaacgaactcttaacatattiaty	1568
Dd	8942	ATTATTTATTTATTTAATTTATTAATTAATATATTTTATTTATTTATTTATTTATATAT	8883
OY	1569	tttagtggaaatttaataagagcggaacttgyaacaatacatattatattagatactagt	1628
Dd	8882	TAAATATATATTTATTTATTTAATAATAA--TTTAAATTTTAAATAATAATAATATAT	8825

OY	1629	atgagattatccaataacacttcttggaagtgtttaacctaaactcgtttccctaagc	1688
Dd	8824	TAAATTTTAAATAAATAATTTCATTTTAATTCATTTAAATAATAATTTTAAATAATTAATTT	8765
OY	1689	cataaatcaatcacgcaggtaaanaaagtttcctatttcgcgatcgagaega	1748
Dd	8764	AATTTATRTATTTTAAACATTTTATTTTAATTTATTTAATATATATATTTTATTTATTA	8705
OY	1749	ttaaccitaaagacttiaatttttggaaaatgtaacccttiacaataaataaacg	1808
Dd	8704	ATTATATATTAATAATTTATATATTTTAAATAATTTAATTTAATTTAATTAATAATTTTAA	8645
OY	1809	catgttttctgcccaataagacaacctcacaactcgtatgctaaattttctgcaa	1868
Dd	8644	TTTTATTAATTAATTAATTAATTTATTTAATATATATATTTTATTTTAAATAATATAT	8585
OY	1869	tattaatitagaacttcaagtcacatcaataagaagaacgcgaglatcacattta	1928
Dd	8584	TATTTATTTTAAATATTTTAAATATTTTATTTAATTTTAAATATATATTTTATTTATTA	8525
OY	1929	atttaagaacaaaaattttgaaaa-----atgtataatttcctacaatatitaaa	1981
Dd	8524	ATTATATTTAATTAATTTTAAATAATAATAGTTATTTATTTAATTAATTAATTAATTAAT	8465
OY	1992	atagaatgcctataatgtaattccatgctctctaaactttttttcataatagtat	2041
Dd	8464	TTATATATTTTAAATAATTTTATTTAAATTTATTAATTAATTAATTAATTTATTTAAATTAAT	8405
OY	2042	aatacatatgaaccaataatagttggtagaattcaaatccatccataataatttga	2101
Dd	8404	TATTTATTTTAAATTTTAAATATTTTATTTAATTTTAAATATATATATATTTTATTTATTT	8345
OY	2102	aactacaattatlaataatttagtcaataacaaatgcatalagaagaatccccaaaaaallt	2161
Dd	8344	ATTATATTTAATTTAATTTAATTAATTAATTTAATTTAATTTAATTTAATTTAATTAATTAAT	8285
OY	2162	tgttaacgagaactccaatttttttttttcttcttggaacaagaataacagatagaac	2221
Dd	8284	TTTATTTTATTTAATTAATTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTAAT	8225
OY	2222	tatttgttctgggaatgaa-gtaagtaataacattaaagcaatttcaaaaattata	2280
Dd	8224	TATTTATTTTAAATTTTAAAATTTTATTTTATTTTATTTTATTTAATTAATTTTATTTATTT	8155
OY	2281	agccatatacgccccaagatgtlctcagtagtggtaattaaatgaatcagctgycgat	2340
Dd	8164	TATTTAATTAATTAATTAATTTTGTTAATTTATTAATTTTAAATATATATTTTATTAATTT	8105
OY	2341	tcaagaattgggaacaacaatgaaaacggaattaaatatataacttaaaataaataa	2400
Dd	8104	TATTTATTTAATTTTATTCATTTCAATTTAATTAATTAATTAATTAATTAATTAATTAATTTCT	8045
OY	2401	ttgagtaaatgtct 2415	
Dd	8044	TTAATTTAATTTATTT 8030	
RESULT	15		
LOCUS	DMU11584/c		
DEFINITION	DMU11584	4601 bp	DNA
ACCESION	U11584		INV
VERSION	U11584.1	GI:508826	
KEYWORDS		mitochondrial DNA; A+T region; tandem repeats.	
SOURCE		fruit fly.	
ORGANISM		Mitochondrion Drosophila melanogaster	
		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
		Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
		Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE		1 (bases 1 to 4601)	
AUTHORS		Lewis,D.L., Parr,C.L., Farguhar,A.L. and Kaguni,L.S.	
TITLE		Sequence, Organization and Evolution of the A+T Region of Drosophila melanogaster Mitochondrial DNA	









GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 : Search time 27.59 Seconds  
(without alignments)  
2460.039 Million cell updates/sec

Title: US-09-502-426-2  
Perfect score: 2681  
Sequence: 1 MFETEHNTLLPLLLPSLILS.....FAFPVDFPNGULPIRYSRL 513

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_16:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2681	100.0	513	10	064989
2	2677	99.9	513	10	064989 arabidops
3	951	35.5	474	10	09SC09
4	866	32.3	512	10	09LKH7
5	864	32.2	524	10	09FX29
6	860	32.1	457	10	023242
7	822	30.7	478	10	09M066
8	815	30.4	482	10	09LW73
9	789	29.4	464	10	09LY89
10	726	27.1	465	10	09LICS
11	711	26.5	465	10	09FMA5
12	658	24.5	463	10	09LH81
13	632	23.6	457	10	09FH76
14	631	23.5	518	10	065624
15	631	23.5	735	10	09F138
16	622	23.2	443	10	09LJ62
17	611	22.8	482	10	09LJK2
18	589	22.0	486	10	081077
19	581	21.7	487	10	09L169
					023384 arabidops

20	581	21.7	496	10	09FOY4	09fy4 cucurbita m
21	579.5	21.6	474	10	09SHY7	09shy7 arabidops
22	579.5	21.6	485	10	09SHY2	09shy2 arabidops
23	577	21.5	455	10	09LKH8	09lkh8 arabidops
24	575.5	21.5	464	10	004949	004949 arabidops
25	569.5	21.2	477	10	09LYV7	09lyv7 arabidops
26	547.5	20.4	489	10	09ZV72	09zv72 arabidops
27	518	19.3	504	10	09SNG3	09sng3 oryza sativ
28	497.5	18.6	460	10	09SYN2	09syn2 arabidops
29	479.5	17.9	444	2	09SYN2	09syn2 arabidops
30	461	17.2	497	11	09R1E4	09r1e4 mus musculu
31	446	16.6	512	4	09NR63	09nr63 homo sapien
32	444.5	16.6	492	13	093323	093323 xenopus lae
33	442.5	16.5	525	4	09NPA1	09npa1 homo sapien
34	439.5	16.4	492	13	09PUB4	09pub4 gallus gall
35	435	16.2	224	10	09SDM6	09sdm6 helianthus
36	398.5	14.9	375	10	09LW32	09lw32 arabidops
37	369	13.8	349	10	09FWR7	09fwr7 arabidops
38	357	13.3	518	5	017807	017807 caenorhabdi
39	338	12.6	453	2	09KFA6	09kfa6 bacillus ha
40	328.5	12.3	517	5	017806	017806 caenorhabdi
41	321	12.0	497	13	09PVI0	09pvi0 fundulus he
42	317	11.8	483	10	09FW33	09fw33 oryza sativ
43	305	11.4	501	11	09QXF7	09qxf7 ratius norv
44	300	11.2	491	11	064463	064463 mus musculu
45	300	11.2	513	10	09SBP8	09sbp8 populus trl

## ALIGNMENTS

## RESULT 1

ID 064989 PRELIMINARY: PRT: 513 AA.

AC 064989; PRELIMINARY: PRT: 513 AA.  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE STEROID 22-ALPHA-HYDROXYLASE.  
GN DWF4.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidops.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WS-2;  
RA Choe S., Dilkes B.P., Fujioka S., Takatsuto S., Sakurai A.,  
RA Feldmann K.A.;  
RL Plant Cell 10:0-0(1998).  
CC -1. SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AF044216; AAC05093.1; -.  
DR InterPro; IPR001128; -.  
DR Pfam; PF00067; P450; 1.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 513 AA: 5867 MW; B1639BDD9A4DA6F3 CRC64;

Query Match 100.0%; Score 2681; DB 10; Length 513;  
Best Local Similarity 100.0%; Pred. No. 6.2e-183;  
Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFETEHNTLLPLLLPSLILSLLPLILKRRKRTRFNLPPKSGMPFGEITGYLKPPT 60  
DB 1 MFETEHNTLLPLLLPSLILSLLPLILKRRKRTRFNLPPKSGMPFGEITGYLKPPT 60  
QY 61 ATTLGDFMOOHYSKYGKIRSNLFCGPTIVSADAGLNRRITLLEGRIFECSPRSIGTL 120  
DB 61 ATTLGDFMOOHYSKYGKIRSNLFCGPTIVSADAGLNRRITLLEGRIFECSPRSIGTL 120  
QY 121 GRWSMLVLVGMDHRDMRSISLNLFLSHARLRTLLKDVRRHTLFLVDSWOONSIFSAODEA 180

```

Db 121 GKMSMLVVGDMHRDMSISLNFLSHARLTKILKDYERHTLFVLDWMQONSIFSADDEA 180
QY 181 KKFENLMAKHMSMDGEETEBOLKKEYVTFMKGVSAPIPLPGTAVHKALOSRAATILK 240
Db 181 KKFENLMAKHMSMDGEETEBOLKKEYVTFMKGVSAPIPLPGTAVHKALOSRAATILK 240
QY 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEMSKSDHYRKORTDDLLGWLKHNSLSTEQ 300
Db 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEMSKSDHYRKORTDDLLGWLKHNSLSTEQ 300
QY 301 ILDLILSLFAGHETSSVAIALAIFFLQACPKAVEELREHELETARAKKEIGESLWMD 360
Db 301 ILDLILSLFAGHETSSVAIALAIFFLQACPKAVEELREHELETARAKKEIGESLWMD 360
QY 361 YKKMDFQCVINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLDNSRYD 420
Db 361 YKKMDFQCVINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLDNSRYD 420
QY 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
Db 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
QY 481 LKENWELAEDDQPPAPFPVDPNGLPTRVSRIL 513
Db 481 LKENWELAEDDQPPAPFPVDPNGLPTRVSRIL 513

RESULT 2
09SC09 PRELIMINARY; PRT: 513 AA.
AC 09SC09;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE (DMF4).
GN T3A5.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bioecker H., Mewes H.W., Lemcke K., Mayer K.F.X., Quetler F.,
RA Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC EMBL: AL132979; CAB62435.1; -.
DR InterPro: IPR001128; -.
DR Pfam: PF00067; P450.1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 513 AA; 58867 MW; B1639BD9A5D7C93 CRC64;

Query Match 99.9%; Score 2677; DB 10; Length 513;
Best Local Similarity 99.8%; Pred. No. 1.2e-182;
Matches 512; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 121 GKMSMLVVGDMHRDMSISLNFLSHARLTKILKDYERHTLFVLDWMQONSIFSADDEA 180
Db 121 GKMSMLVVGDMHRDMSISLNFLSHARLTKILKDYERHTLFVLDWMQONSIFSADDEA 180
QY 181 KKFENLMAKHMSMDGEETEBOLKKEYVTFMKGVSAPIPLPGTAVHKALOSRAATILK 240
Db 181 KKFENLMAKHMSMDGEETEBOLKKEYVTFMKGVSAPIPLPGTAVHKALOSRAATILK 240
QY 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEMSKSDHYRKORTDDLLGWLKHNSLSTEQ 300
Db 241 FIEKMEERKLDIKEEDQEEBEVTEDEAEMSKSDHYRKORTDDLLGWLKHNSLSTEQ 300
QY 301 ILDLILSLFAGHETSSVAIALAIFFLQACPKAVEELREHELETARAKKEIGESLWMD 360
Db 301 ILDLILSLFAGHETSSVAIALAIFFLQACPKAVEELREHELETARAKKEIGESLWMD 360
QY 361 YKKMDFQCVINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLDNSRYD 420
Db 361 YKKMDFQCVINETLRIGNVYRFLHRKALDVRKGYDIPSGMKVLPVISVHLDNSRYD 420
QY 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
Db 421 QPNLFNPMWQOONNGASSGSGSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHHLV 480
QY 481 LKENWELAEDDQPPAPFPVDPNGLPTRVSRIL 513
Db 481 LKENWELAEDDQPPAPFPVDPNGLPTRVSRIL 513

RESULT 3
09LKH7 PRELIMINARY; PRT: 474 AA.
AC 09LKH7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CYTOCHROME P450.
GN CYP7.
OS Phaseolus aureus (Mung bean) (Vigna radiata).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Vigna.
OX NCBI_TaxID=3916;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. 2937;
RA Yang M.T., Chen Y.M.;
RT "Cloning and sequencing of a Vigna radiata cDNA encoding cytochrome
RT P450.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY
CC SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AF279252; AAF89209.1; -.
DR InterPro: IPR001128; -.
DR Pfam: PF00067; P450.2.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Electron transport; Endoplasmic reticulum; Heme; Membrane; Microsome;
KW Monooxygenase; Oxidoreductase.
SQ SEQUENCE 474 AA; 54037 MW; A2542A809C5BAC6D CRC64;

Query Match 35.5%; Score 953; DB 10; Length 474;
Best Local Similarity 40.1%; Pred. No. 8.2e-60;
Matches 203; Conservative 97; Mismatches 162; Indels 44; Gaps 10;

```

```

QY 1 MFEETHHTLPLLLPSLLSLFLLLKRRNRKTRFNLPKSGMFPGLGTTGYLKPPT 60
Db 1 MFEETHHTLPLLLPSLLSLFLLLKRRNRKTRFNLPKSGMFPGLGTTGYLKPPT 60
QY 61 ATTLDGDMOQHVSKYKGRIVSNLGEPTIVSADAGLNRFILQNGRLFECSYPSRSTGIL 120
Db 61 ATTLDGDMOQHVSKYKGRIVSNLGEPTIVSADAGLNRFILQNGRLFECSYPSRSTGIL 120

```

```

QY 12 LLLPSLLSLF-----LLLRNRNRKTRFNLPKSGMFPGLGTTGYLKPPTATTG 66
Db 1 MVSLEPLT--LLTFAASAAATFLHRAFSRRKFRLLPGSYGLPFGITQLLSAVKSSNPEP 58

```

```

QY 67 FMOQHSKYKTYRSLNLFGEPTIVSADAGLNRFILONEGRLEFECSPYRSIGGLGKWSML 126
DB 59 FMDENVRKRGSLFNKTHVEBEPTVFSADPELNRFILONEGRKLDSCSPGISNLLGKHSLL 118
QY 127 VLVGMHMDRMSISLNFSLHARLRTLLKADVERHFLVLDSSWOQNSIFSADAEKAKTEPN 186
DB 119 LKKGALHKRMSHSLTMSFANSIILKDLHHIDRLIGLNDJWSDR--VYLMOAKKITPE 176
QY 187 LMAKHMSDPEEETEOLKKEVYPMKGVASAPLNPCTAHKALQSAATILKFERKM 246
DB 177 LVYKQMSDLP--DEWTSRKKEVLEBGFPLPLPSTYRRARAKATKVAEALTIV 235
QY 247 EERKLDIKEDEDEEVEKTEDEAEMSKSDHVRKORTDLDGLGWL--KHSNLSLEQILDI 305
DB 236 RQR---REEVNOKKEKS-----DMGALLASGSHFDDQIVDL 272
QY 306 LSLFRAGHETSVAIALAIFLQACPKAVEELREHLEIARAKKEIGESLWMDYKKKD 365
DB 273 LALLVAGYETSTIMTLAVKFLTEPLPLALQKEHDOJ--RARSDFG--APLEMTWYKSKV 330
QY 366 FPOCVINETRLGNVYRFLHRKALKDVRKYGDIPSGMVLVISAIVHLDNSRYDQPNLF 425
DB 331 FTOHYVNETLRANITIGGIFRATTDIDIKGTTIRGMVVERFASFRVHNLNPEYKDATFE 390
QY 426 NPMRMOQONGASSSGSGSFTWGNMYPFGGPRCLAGSELAKLEMAVFIHLVLKFNW 485
DB 391 NPMRMOQNSSEANP-----ANYTTPFGGPRCLCPGYELARVLSVFLRIYVRFSGM 442
QY 486 ELAEDDOPPAFPVDFPNGLPIRVSR 511
DB 443 VPAAEDKLVFPPTTRQKRYPIIVKR 468

```

```

RESULT 4
Q9FX29 PRELIMINARY: PRT: 512 AA.
AC Q9FX29;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE STEROID 22-ALPHA-HYDROXYLASE, PUTATIVE.
GN T9L24.44.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RA "Arabidopsis thaliana chromosome 1 BAC T9L24 genomic sequence.";
RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
RL EMBL; AC012396; AAG30983.1; -.
DR DR
SQ SEQUENCE 512 AA: 58297 MW: 1590C0304BF319FA CRC64;

```

Query Match 32.3%; Score 866; DB 10; Length 512;  
 Best Local Similarity 33.8%; Pred. No. 1e-53;  
 Matches 181; Conservative 105; Mismatches 173; Indels 76; Gaps 7;

```

QY 12 LLLPSLSLLLELTL-----LKRNRKTRFNLPPCKSGWPLGELTIGLKYPTATLTLD 66
DB 11 LLSVSSSTFLFAITIFLLAGIARRRRARPHRLPPSGSRWPLIGDFALNVAAGSHPS 70
QY 67 FMOQHSK-----YGRYSNLFGEPTIVSADAGLNRFIL 101
DB 71 FYEKQIKRKFVSLCVALILKRPDNGSFRNRYGRIFGSCSLFGKAAVYADPDFKRFIM 130
QY 102 ONEGRLEFECSPYRSIGGLGKWSMLVGDHMDRMSISLNFSLHARLRTLLKADVERHT 161

```

```

DB 131 ONEGRLEFSSYPKSPFRDLVCGKXGVITVHGDDQQRLLHSIASSMMRHDLKTHFLVPIVYM 190
QY 162 LFLVLDSSWOQNSIFSADAEKAKTEPNLMAKHMSMDPEEETEOLKKEYTFRKGVASAPL 221
DB 191 LOTLSNFKDGEVVLLODICKRVAIHLMNQILGVS--SESEVDEMSQLFSDPFGCLSVPI 249
QY 222 NLPCTAHKALQSAATILKFERKMDKEEDQEEVEKTEDEAEMSKSDHVRKQR 281
DB 250 DLPGFYNNAMAKKEITIKKTIKRLONKAASD-----T 286
QY 282 TDDDLGLWVLKHSNSTEOILDLISLFRAGHETSVAIALAIFLQACPKAVEELREH 341
DB 287 AGNGVGLRLEESLPNESMADEFTINLFRAGNETTSKTLFAYVFLHCPRKMTQLEEH 346
QY 342 LEIARAKKEIGESLWMDYKKADFTQCVINETRLGNVYRFLHRKALKDVRKYGDIPS 401
DB 347 -----DRLAGGMLTWQDKTMDFTQCVIDETLRLGGIAIWLMPRAKDVSYODVPIRK 399
QY 402 GKWVLVPISAVHLDNSRYDQPNLFPWRV-----OQONGASSSGSGSFTWGNMYPFG 456
DB 400 GCFVVPFLSAVHLDESYYEESLSPNWRMLDPETOQKRMRKTSF-----FYCPFG 449
QY 457 GGPRLCAGSELAKLEMAVFIHLVLKFNWELAEEDQPFAPVDFPNGLPIRVSR 511
DB 450 GGTRECPGAEIARLQIALFLHYFITTYKWTQLEKDRISFPSPARLVNGFKIQLNR 504

```

```

RESULT 5
O23242 PRELIMINARY: PRT: 524 AA.
AC O23242;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-AUG-1999 (TREMBlrel. 11, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CYTOCHROME P450 (CYTOCHROME P450 LIKE PROTEIN).
GN ROTUNDIFOLIA3 OR C7A10.980.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons: core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361880; PubMed=9694802;
RA Kim G., Tsukaya H., Uchimiya H.;
RT "The ROTUNDIFOLIA3 gene of Arabidopsis thaliana encodes a new member
RT of the cytochrome P-450 family that is required for the regulated
RT polar elongation of leaf cells.";
RL Genes Dev. 12:2381-2391(1998).
RN [2]
RP SEQUENCE OF 68-524 FROM N.A.
RA Bevan M., Terry N., Vos P., Heijnen L., Mewes H.W., Schueller C.,
RA Chaiwatzis N.;
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE OF 270-524 FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Hohnsels J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.,
RA Schueller C.;
RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOOPROTEIN + O(2) = ROH +
CC OXIDIZED FLAVOOPROTEIN + H(2)O (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DB EMBL; AB008097; BAA37167.1; -.
DB EMBL; Z99708; CAB16850.1; -.
DB EMBL; AL022141; CAA18139.1; -.
DR InterPro: IPR001005; -.
DR InterPro: IPR001128; -.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.

```

DR PROSITE: PS00037; MYB\_1; UNKNOWN\_1.  
 KW Electron transport; Endoplasmic reticulum; Heme; Hypothetical protein;  
 KM Membrane; Microsome; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 524 AA; 59389 MW; 550578908BDDF272 CRC64;

Query Match 32.2%; Score 864; DB 10; Length 524;  
 Best Local Similarity 35.4%; Pred. No. 1.5e-53;  
 Matches 186; Conservative 96; Mismatches 165; Indels 78; Gaps 10;

24 FLII-----LKRNRKTRFN-----LPPKSGMPFGETIGYL--- 56  
 30 FLVLPAGILLRPMIMFRMLRNKSTKGDDEEDNEEKGMIPNGSLGWPFGETLNFACG 89  
 57 ---KPYATTLGDPMOQHVSKYKIRSNLGEPTIVSADAGLNRFLQNGRLFECSYP 113  
 90 YSSRPV-----FMDKRKSLYGVKFTNIIIGTPIIISTDAEVKAYVLQNGNFFVAPY 143  
 114 RSIGGILGKMSMLVVGDMHDMRSISLNTLSHARLTIILIKDYERHTLVLDSSMOONSI 173  
 144 KSTIELLENSILSINGPHOKRLHTLIGAFLRSPHLKDRITRDIASVVLTLASMAOLPL 203  
 174 FSADAEAKKFFENIMAKHIMSMDPEETEQLKKEVYFMKGVVASPLNPGTAYHKALQ 233  
 204 VHVDEIKKMTFEILVKYLMSTSPG-EDMNLILKEFEFGLICIPKFGTLYKSLK 262  
 234 SRATILFERKMEERKLDIKEEDQEEVEYK-----EDEAKMSDHYAKQRTDIDL 287  
 263 AKERLIKVKVVEERQVAMTTTSPANDVYVLLRDGDSKQSPDSFGSK----- 315  
 288 GWVLKHSNLSTEQILDLISLTFAGHETSSVAIALAIFLQACRAVELREHELETARA 347  
 316 -----IVEMMIPGEETMTAMTLAVKFLSDNPALAKLVEENMEKRR 358  
 348 KKEIGESLENDYKKMDTCVINEETRLGNVRFELRKAKDKVRYKGYDIPSGMKVLP 407  
 359 KLELGE-EYKKTDMSLSTFQNVINETLRMANIINGWRKALKDVEIKGYILPQWCYLA 417  
 408 VISAHVLDNSRYDQPNLFNPRMOQONNGASSGSGSESTWGNVMPFGGPRLCAGSEL 467  
 418 SFISVHMEDEIDYDPQFDPWRMDRINGSANSSIC-----FTFPGGQRLCPGLEL 468  
 468 AKLEMAVFHHLVLEKFMNELAEDDQFAFPVDFPNCGLPIVSRIL 512  
 469 SKLEISTFLHLVTRYSM-TAEDEIVSFTVKKMRRLPIRVATV 512

RESULT 6  
 09M066 PRELIMINARY; PRT; 457 AA.  
 AC 09M066;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE CYTOCHROME P450 LIKE PROTEIN.  
 GN ATAG36380.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 ON NCBI\_TaxID=3702;  
 RX 11  
 RP PARTIAL SEQUENCE FROM N.A.  
 RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,  
 RA Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RA Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
 -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY

CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL: AL161589; CAB80304.1; -  
 DR InterPro: IPR001005; -  
 DR InterPro: IPR001128; -  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PRO0385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 DR PROSITE: PS00037; MYB\_1; UNKNOWN\_1.  
 KW Electron transport; Endoplasmic reticulum; Heme; Membrane; Microsome;  
 KM Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 457 AA; 51607 MW; BD6A148E44EACD2 CRC64;

Query Match 32.1%; Score 860; DB 10; Length 457;  
 Best Local Similarity 36.8%; Pred. No. 2.4e-53;  
 Matches 179; Conservative 93; Mismatches 160; Indels 54; Gaps 8;

39 LPPKSGMPFGETIGYL-----KPYATTLGDPMOQHVSKYKIRSNLGEPTIVSA 92  
 2 IPNGSLGMPVIGETLNFACGYSSRPVT-----FMDKRKSLYGVKFTNIIIGTPIIIST 55  
 93 DAGNRFILQNGRLFECSYPRSTIGILGKMSMLVVGDMHDMRSISLNTLSHARLRTI 152  
 56 DAEVKKVYVLQNHGNTFVPAYPKSTIELLENSILSINGPHOKRLHTLIGAFLRSPHLKDR 115  
 153 LKDYERHTLVLDSSMOONSIIFSADAEAKKFFENIMAKHIMSMDPEETEQLKKEVYF 212  
 116 ITRDIASVVLTLASWQMLPVHQDEIKKMTFEILVKYLMSTSPG-EDMNLILKEFEF 174  
 213 MKGVASAPLNPGRYAHKALQSRATILKFERKMEERKLDIKEEDQEEVEYK-----E 266  
 175 IKGLICIPKFGRTLYKSLAKERLIKVKVVEERQVAMTTTSPANDVYVLLRDGSD 234  
 267 DEAKMSDHYAKQRTDIDLGVWLKHSNLSTEQILDLISLTFAGHETSSVAIALAIF 326  
 235 SEKQSPDSFGSK-----IVEMMIPGEETMTAMTLAVK 270  
 327 LQACRAVEELREHELETARA KKEIGESLENDYKKMDTCVINEETRLGNVRFELR 386  
 271 LSDNPVALAKLVEENMEKRRKLELGE-EYKKTDMSLSTFQNVINETLRMANIINGWR 329  
 387 KALNDVRYKGYDIPSGMKVLPVISAHVLDNSRYDQPNLFNPRMOQONNGASSGSGSFS 446  
 330 KALNDVEIKGYLIRKGCYVLA SFTSVHMEDEIDYDPQFDPWRMDRINGSANSSIC---- 385  
 447 TWGNVMPFGGPRLCAGSELALEMAVFIHHLVLEKFMNELAEDDQFAFPVDFPNCGLP 506  
 386 -----FTFPGGQRLCPGLELSKLEISTFLHLVTRYSM-TAEDEIVSFTVKKMRRLP 439  
 507 IRVSRIL 512  
 440 IRVATV 445

RESULT 7  
 09LNT3 PRELIMINARY; PRT; 478 AA.  
 AC 09LNT3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE T12C24.27.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 ON NCBI\_TaxID=3702;  
 RX 11  
 RP SEQUENCE FROM N.A.  
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,  
 RA Shin P., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,

RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
RA Ecker J.R.;  
"Genomic sequence for Arabidopsis thaliana BAC T12C24 from chromosome  
1";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chlou J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL: AC025417; AAF8087.1; -  
DR InterPro: IPR001128; -  
DR Pfam: PF00067; p450; 1.  
DR PRINTS: PR00385; P450.  
DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 478 AA; 55057 MW; F6A9B9F2FD8558 CRC64;

Query Match 30.7%; Score 822; DB 10; Length 478;  
Best Local Similarity 35.6%; Pred. No. 1.3e-50;  
Matches 181; Conservative 95; Mismatches 187; Indels 46; Gaps 8;

QY 12 LLLPSLLSLFLLLKRRNKTFRNLPKSGWPFLETGTGYLAKPTATTLGDFMOOH 71  
DB 4 LLIWVSLIISTHWWYSWRNPKCRGKLPKSGMFPPLGEGSIOFPKPTSDIPPIKER 63  
QY 72 VSKYKIRSNLFGEPTIYASADAGNRFILONEGRLEFECYSYRSGIGIKKSMVLVD 131  
DB 64 VKKYPIRTKTVGRPVIVSTDADSLYFVNOEGRCFQSWYDTTTHIFGKNNVSLHGF 123  
QY 132 MHRDMSISLFLSHARLRTLLKDYERHTLVLDSDWOONSIFSAODEAKKFTMLMAKH 191  
DB 124 MYKYLKNNVLTFLGHDLKK-MLPQVEMTANKRLLELMSNODSVELKDATASIFDLTKK 182  
QY 192 IMSMPGEBETQOLKKEYTFMKGVASAPLNPSTAYHKALOSRTILKIFRKKEBERL 251  
DB 183 LISHDP-KKSSENLRANVAFIOGLISPEFDIPGTAYHNCLOGRAKAKMLFNMLOERR- 240  
QY 252 DIKEDDOEEBEVKTTDEAEMSKSDHYVRQRTD--DDLGGVLLKHSNLSTEOI-LDLIISL 308  
DB 241 -----ENPRKNPSDFPDYVIEIOKEGTILTEELALDLMEVL 277  
QY 309 LEAGHETSSVAIALAIFLOACPKAVEELREEHLEIARAKKELGSELNMDYKKMDFTO 368  
DB 278 LEASETTSIALTAIKFLSDDEVLKRLTEHEHTILR-NRSDADSGULWEYKSMITYYF 336  
QY 369 CVINETLRLGNVRLHRRALKDVRK-----GYDPSGKVLVPLVISAHLNDSRYDOP 422  
DB 337 OFINETARLANIVPAIFRKAFLDIKEFEVNDTYTIPAGNVMVNCPPRAVHNLNPEYKXP 396  
QY 423 NLFNPMRMOOQNNGASSSGSFSFTWGNVMPGGGPRLCAGSELAKEMAVFIHILYLK 482  
DB 397 LVFNPSRME-----GSKVTNASKHFMALFGGMRFCVGTDTTKLQMAFLHSLVTK 446  
QY 483 FNNLEADDDQPAFPFVDFPNCGLPIRVSR 511

DB 447 YRMEIKGNITRTPLGLOFPNGYHVKLK 475  
RESULT 8  
QYLY89 PRELIMINARY; PRT; 382 AA.  
AC QYLY89;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)  
DE HYPOTHETICAL 43.9 KDA PROTEIN.  
GN F18022-190.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN  
RP SEQUENCE FROM N.A.  
RA Beyan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,  
RA Rudd S., Lemke K., Mayer K.F.X.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY  
CC SIMILARITY).  
DR EMBL: AL163817; CAB87779.1; -  
DR InterPro: IPR001128; -  
DR Pfam: PF00067; p450; 1.  
DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Electron transport; Endoplasmic reticulum; Hypothetical protein;  
KW Membrane; Microsome.  
SQ SEQUENCE 382 AA; 43888 MW; 1BC5685AB7384E30 CRC64;

Query Match 30.4%; Score 815; DB 10; Length 382;  
Best Local Similarity 34.2%; Pred. No. 2.9e-50;  
Matches 159; Conservative 87; Mismatches 127; Indels 92; Gaps 4;

QY 45 GMPFLETGTGYLAKPTATTLGDFMOOHVSKYKIRSNLFGEPTIYASADAGNRFILONE 104  
DB 2 GMPFLETGTISFEKPRHSDSIGFTLQORVSRIGKVKRSNIGCKAVASCDQELMFTLONE 61  
QY 105 GRLEFECYSYRSGIGIKKSMVLVDGMHRDMSISLFLSHARLRTLLKDYERHTLV 164  
DB 62 GKLFTSDYPRKAMHDLGKYSLLTAGETIRKLKLVNIIISFINLTKSPDLHCAENLISLI 121  
QY 165 LDSWOONSIFSAODEAKKFTFMMAKHNSMDPGEDETEQOLKKEYTFMKGVASAPLNP 224  
DB 122 LKSMNCKREVERHKEVKIFTTLSVMNQULSIKPREPARLYVLVDLSYKGFISPIPLP 181  
QY 225 GTAAYHKALOSRTILKIFERKKEBERLDIKEDDOEEBEVKTTDEAEMSKSDHYVRQRTD 284  
DB 182 GTGYTNAIKVRNRMNINHONAITEDMNNALREDFLDISTISNED----- 224  
QY 285 DLLGGVLLKHSNLSTEOIDLILSLFAGHETSSVAIALAIFLOACPKAVEELREEHLEI 344  
DB 225 -----EEHNAI 230  
QY 345 ARAKKELGSELNMDYKKMDFTQCVINETLRLGNVRLHRRALKDVRKGYDIDPSGK 404  
DB 231 -RAKKAGDEL-LNWEYQKMEFTQCVISALRGNIIVKTHHKAHNDIKFNENYVIPDKWK 288  
QY 405 VLPVISAHLNDSRYDOPRLFNPWRMOOQNNGASSSGSFSFTWGNVMPGGGPRLCAG 464  
DB 289 VPIFTFAHLSDSLHENPEENPMKMTKTT-----AFGGGVAVCPG 329  
QY 465 SELAKEMAVFIHILYLKFNWELAEDDQPAFPFVDFPNCGLPIRV 509

Db 330 GELGKLDIAFLHLVLSYRKIKSDENPIAHVPERKMLLEI 374

RESULT 9

09LICS PRELIMINARY: PRT: 464 AA.

AC 09LICS: 01-OCT-2000 (Tremblrel. 15, Created)

DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE CYTOCHROME P450-LIKE PROTEIN.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA:

RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA:

RX PubMed=10907853;

RA Nakamura Y.;

RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety P1, TAC and BAC clones."

RL DNA Res. 7:217-221(2000).

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL: AP001307; BAB01922.1; -.

DR InterPro: IPR000504; -.

DR InterPro: IPR001128; -.

DR Pfam: PF00067; P450; 1.

DR PRINTS: PR00385; P450.

DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.

DR Heme; Monooxygenase; Oxidoreductase.

KW SEQUENCE 464 AA; 52864 MW; A65E09465E5190B CRC64;

Query Match 29.4%; Score 789; DB 10; Length 464;

Best Local Similarity 36.5%; Pred. No. 2.7e-48;

Matches 163; Conservative 101; Mismatches 145; Indels 38; Gaps 8;

Db 40 PPKSGMPFLEGTGYLKPYATTLGDMOAHVSKGIYRSNLFGEPITVSADAGLNR 99

Db 52 PHSILGMPVIGETIEFVSAYSADRESFMDKRLMYGVRFSHIFGTATVSTDAEVNRA 111

Db 100 ILONEGRLEFESYPRISIGILGKMSLVLDGMDHMRKRSISLNFSLARLRTILKQVER 159

Db 112 VLQSDSTAFAFVPEYKTYRELMGKSSILLINGSLHRRHGLVGLKSPDLAKQIVRDMHK 171

Db 160 HTLVLDLSMOONSIFSAODEAKKFTFNILMAKHMSMDPGEETDLKKEVYTFPKGVSA 219

Db 172 FLSSMDLMSDQVLLDQVSKYAFKVLAKALISVEKG-EDLEELKREPENFISGLMSL 230

Db 220 PLNLPGTAHYKALQSRATILKFERKMEERKLDIKEEDQEEVEVTEDEAM--SKSD 276

Db 231 PINPGTQLHRSIQAKKNMVKQYERIEGKIR--KTKNKEEDVIADVDVLLKDSSEH 288

Db 277 VRKORTDDDLGWLKSNLSTEDILILSLFAGHETSSVAIALAIFLQACPKAVEE 336

Db 289 -----LTH-NLIANNMIDMMI-----PGHDSVAVLITLLVKSLSPPALNLT 329

Db 337 LREELHETAKKKESELNMDQYKKMDFTQCYINFTLRLGNVVRHLHRAKALQDVYKG 396

Db 330 LTEENMKL-RSLKELTEBPYLVNDYLSLPFOKAVITFLRGNVITGVMRKAMADVIEKG 388

Db 397 YDIPSGKVLPLVSAVHLNDSRYDQPLNFMWMOQONNGSSGSGSFSTWCNNYMPFG 456

Db 389 YVDPKMGCFPLAVLSVHLDKLKYESPYPKPNMRMRQERDMNTSS-----FSPFG 436

QY 457 GGPRLCAGSELAKLEMAVFTHHLVLPK 483

Db 437 GGRRLCPGLDLARLETSVFLHVLVTRF 463

RESULT 10

09FMA5 PRELIMINARY: PRT: 465 AA.

AC 09FMA5: 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)

DE CYTOCHROME P450.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-COLUMBIA:

RX MEDLINE=98290546; PubMed=9628582;

RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N., Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV. Sequence features of the regions of 1,456,315 bp covered by nineteen P1 and TAC clones."

RL DNA Res. 5:41-54(1998).

DR EMBL: AB009048; BAB08653.1; -.

SQ SEQUENCE 465 AA; 53767 MW; 0C00459C9C86601F CRC64;

Query Match 27.1%; Score 726; DB 10; Length 465;

Best Local Similarity 32.3%; Pred. No. 8.4e-44;

Matches 166; Conservative 91; Mismatches 181; Indels 76; Gaps 10;

Db 12 LLLPLSLLLFLILIKRN--RKTRENLPPKSGMPFLEGTGYLKPYATTLGDFMQ 69

Db 11 LLIVSLCSALL-----RNNOMRYTKNGLPKPGMGPIGETTEFLKQSP-----NEMR 59

Db 70 QHVSQYKIRYSNLFGEPTVSADAGLNRFLIOEGRLFGCSYPRISIGILGKMSMLV 129

Db 60 NQRLRYTSFKSHLGGPPTLSMDSEVNRITLKESGLVPGIPQSLDLGTCNMAVH 119

Db 130 GDMHRMRSISLNFSLARLRTILKQVERHTLVLDLSMOONSIFSAODEAKKFTF---- 185

Db 120 GSHRLKRGSLSLISSYMRDHLRPVDFPMRSYLDQWMELEVTDIDQTKHMAFLSL 179

Db 186 -----NLMAKHMSMDPGEETDLKKEVYTFPKGVSAFLNLPGTAYHKAQSRATILK 240

Db 180 TQINGNLRKRPV-----EEFKTAFPKLVAGTSLVPIIDLPGTNYRCGIARNNIDR 229

Db 241 FIERKMEERKLDIKEEDQEEVEVTEDEAMSKSDHVRKORTDDDLGWLKSNL-----L 296

Db 230 LRLKMEER-----DSGE-----TFDMLGILMKKEGNNPFL 262

Db 297 STEQIIDLILSLFAGHETSSVAIALAIFLQACPKAVEELREELHETAKKKEGSEL 356

Db 263 TDEIRDQVVTIYSGYETVSTSMALKYLDHPKALQELRAENLAFRRKRQ--DEPL 320

Db 357 NMWDYKKMDFTQCYINFTLRLGNVVRFLHRAKALQDVYKGYDIPSGKVLPLVSAVHLN 416

Db 321 GLEDVKSKMFTTRAVIYETSLRATIVNGVLRKTRDELINQVLLIPKGRIRIVYTRRIYDA 380

Db 417 SRYDQPLNFMWMOQONNGSSGSGSFSTWCNNYMPFGGPRLCAGSELAKLEMAVFI 476

Db 381 NLIEDPLIFNPMRMKKSLSQ-----NSCIVFGGRLCPGKELGIVEISSFL 429

Db 477 HHLVLEKFWELAEDDQPFAPFVDFPGLPIRVS 510

Db 430 HYFVTRRMWELIGDDELMAVFPFVAFAPKGFHLRIS 463

RESULT 11  
ID 09LH81 PRELIMINARY: PRT: 465 AA.  
AC 09LH81.  
DT 01-OCT-2000 (TREMBLREL. 15, Created)  
DT 01-OCT-2000 (TREMBLREL. 15, Last sequence update)  
DT 01-MAR-2001 (TREMBLREL. 16, Last annotation update)  
DE CYTOCHROME P450.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA Kaneo T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX PubMed=10907853;  
RA Nakamura Y.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety P1, TAC and BAC clones."  
RL DNA Res. 7:217-221(2000).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AP002060; BAB02270.1; -  
DR InterPro; IPR001128; -  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 465 AA: 53862 MW; 630A21D0765ED0D CRC64;

Query Match 26.5%; Score 711; DB 10; Length 465;  
Best Local Similarity 32.1%; Pred. NO.9.8e-43;  
Matches 163; Conservative 93; Mismatches 195; Indels 56; Gaps 8;  
QY 14 LPLSLTLFLTLK-----RRNRKTRFNLPGKSGMPFLGTTGYLKPYATTLGDF 67  
DB 3 IMMILGLVITVCTALRRNQMRYSKGLPPTGMGPRTGTTETFLKGP-----DF 57  
QY 68 MOQVSKYTGKIRSNLFGPTIVSADAGLNRFILQNGRLFECSYPRISIGILGKMSLV 127  
DB 58 MNQRLRYGSEFKSHILGCPITVSMQAEINRYILMESKGLVAGYPOSMLDILGTCNIA 117  
QY 128 LVGDHHRQRSISLNFSLHARLTLLKDYERHTLFVLDSMOONSIFSQODEAKKFTPL 187  
DB 118 VGPFSHRLMRGSLSLISTPMKDHLPKIDPMRNYLGGWDLDTVDIOETKHNMF-L 176  
QY 188 MAKHISMDPGEETEQLEKEVTFMKGVASAPLNPGTAVYHKALOSRAITLFIERKME 247  
DB 177 SSLQIAETLKRPVEBEYETFEFKLVGLSVPIIDPGINYSQGARANNIDRLLELMQ 236  
QY 248 EKKLDKEEDQEEBEYKTEDEAEMSKSDHVRKORTDDLLGWLKHSN---LSTEQILD 303  
DB 237 ERK-----ESGETFT-----DMGLYLMKKEDNRRLTLDEIRD 269  
QY 304 LTLSLFAGHETSVAIALAIFLQACPKAVEIREHEHLEIRAKKELGESELMNDYRK 363  
DB 270 QVVTILYSYETVSTSMALXYLHDHPALBELREHLEIRKRP--DEPLTLDDIKS 327  
QY 364 MDTQVINEITLGNVYFELHRAKLDVRYKGYDIPSGMKVLPVISAHLDSRYDQPN 423  
DB 328 MKETRAVIEPETSRIATVINGVLAKTHDELNGYLIPKGMRIYVYREINYSIATEDPM 387  
QY 424 LNPFRMOQONNGASSGSGSFSTGCMNMYPRGGCPRLCAGSELAKEAMVFIHLVLF 483  
DB 388 INPFRMMEKSLDESKS-----YFLFGGVRLCPGKELGISSEVSLHYEVTKY 436

QY 484 NMELAEDDQFAFPVDPENGLPIRVS 510  
DB 437 KRENGEDKLMVFPVSAKPGYHICKS 463  
RESULT 12  
ID 09FH76 PRELIMINARY: PRT: 463 AA.  
AC 09FH76.  
DT 01-MAR-2001 (TREMBLREL. 16, Created)  
DT 01-MAR-2001 (TREMBLREL. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLREL. 16, Last annotation update)  
DE CYTOCHROME P450.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
OC Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE=20181125; PubMed=10718197;  
RA Sato S., Nakamura Y., Kaneo T., Kato T., Asamizu E., Kotani H.,  
RL DNA Res. 7:31-63(2000).  
DR EMBL; AB020744; BAB10255.1; -  
SQ SEQUENCE 463 AA: 52366 MW; CDD17293F553F812 CRC64;

Query Match 24.5%; Score 658; DB 10; Length 463;  
Best Local Similarity 31.8%; Pred. NO.5.8e-39;  
Matches 163; Conservative 95; Mismatches 184; Indels 70; Gaps 15;  
QY 14 LPLSLTLFLTLK-----RRNRKTRFNLPGKSGMPFLGTTGYLKPYATTLGDF 68  
DB 6 LPLTSLAALFLCLLFLTAGVRRSSSTKLPLPPTGWTGYVET--POLYSQDP-NVFF 61  
QY 69 QOHVSKYTGKIRSNLFGPTIVSADAGLNRFILQNGRLFECSYPRISIGILGKMSLV 128  
DB 62 AAKQRYGSEFKSHILGCPITVSMQAEINRYILMESKGLVAGYPOSMLDILGTCNIA 121  
QY 129 VGDHHRQRSISLNFSLHARLTLLKDYERHTLFVLDSMOONSIFSAODEAKKFTPL 188  
DB 122 QGDYHSHRLMRGSLSLISTPMKDHLPKIDPMRNYLGGWDLDTVDIOETKHNMF-L 179  
QY 189 AKHISMDPGEET---EQLEKEVTFMKGVASAPLNPGTAVYHKALOSRAITLFIERKME 245  
DB 180 LISL---GKDEVYRREDLKRCYILEKGYNSMPLNLTGTLFHRAMKARKELAOLIANI 235  
QY 246 MEERKLDKEEDQEEBEYKTEDEAEMSKSDHVRKORTDDLLGWLK-HSNLSTEQILD 304  
DB 236 LSKRR-----QNPSSNR-----DLGSMEDKAGLTLDQIAN 268  
QY 305 LTLSLFAGHETSVAIALAIFLQACPKAVEIREHEHLEIRAKKELGESELMNDYRK 364  
DB 269 IIGVIFAARDTASVLTWILKYLADNPVLEAVTEQMAIRDKKE-GES-LTWEDTKKM 326  
QY 365 DFTQVINEITLGNVYFELHRAKLDVRYKGYDIPSGMKVLPVISAHLDSRYDQPN 424  
DB 327 PLTYRIQETLRAATLSTFEREADVEYEGLLIPGKMKVLPLEFNHNDISDPQPK 386  
QY 425 FNPFRMOQONNGASSGSGSFSTGCMNMYPRGGCPRLCAGSELAKEAMVFIHLVLF 484  
DB 387 FDPSPRE-----VAPKPTFMFPGSGHISCGNELAKLEISVLVHLHTTKR 433  
QY 485 WELAEDDQ-----PFAFPVDPENGLPIRVS 511  
DB 434 WSIYVPSDGIQYCPALP-----QNGPLIALER 461



```

RESULT 13
ID 065624 PRELIMINARY; PRT; 457 AA.
AC 065624;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE CYTOCHROME P450.
GN T18B16.200 OR A74G19230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Beves V., Rechmann S., Borkova D., Ansoorge W., Bancroft I.,
RA Mewes H.W., Mayer K., Scheller C.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Beves V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 131-457 FROM N.A.
RA Van Der Schueren J., Chuang Y.J., Aert R., Defoor E., Robben J.,
RA Vockert G., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AL021687; CA116713.1;
DR EMBL: AL161550; CAB78925.1;
DR HSP; P33006; ICP1.
DR InterPro: IPR001128;
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; UNKNOWN_1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ
SEQUENCE 457 AA; 52436 MW; 6105FB7C181E4F07 CRC64;

Query Match 23.6%; Score 632; DB 10; Length 457;
Best Local Similarity 31.8%; Pred. No. 4e-37;
Matches 151; Conservative 92; Mismatches 176; Indels 56; Gaps 12;

QY 13 LLDLSLILFLILKRRNKRTPNLPKSGMPFLGETIGYLPYATTLGDPMOOHV 72
DB 10 LFAQSLFYLRCLISORFGSSKILPPTGMPYVGET---POLYSODP-NVFFQSKQ 65
QY 73 SKYKTVRSNLFGEPTIYSADAGLNRFILONEGLFECSYRSTIGILGKMSMLVLVGM 132
DB 66 KRYSSVFETHVLCGPCVWISSPEAKFVLYTKSHLFKTFPAKSKRMKGLKATFFHOCGY 125
QY 133 HRDKSISLNLFLSHARLTILKDVERTTLVLDWMQONSIFSAODEAKKTFNIMAKHI 192
DB 126 HAKLRKLVLRAPFPESTIRN-WVPDIESIAODLSRW-EGTMINITYQEKTYTFVALLSI 183
QY 193 MSMDGEEET---EOLKEVYTFMKGVYAPLNPATYHAKALOSRATILKFIERKEER 249
DB 184 F---GKDEVLYRDLKCYLYLEKGYNMPVNLPGTLFHSKMRKRLSOLILARIISER 239
QY 250 KLIDKEEDQEEEVKTEDEAKMSKSDHVRKORTDDLLG-VLKLHNSLSTEQIIDLITSL 308
DB 240 R-----ONGSSR-----NDLGSFGKKEELTDQIDADNITIGV 272
QY 309 LFAHETSSVAIALAIFLQCPRAVELREHLEIARAKKELESELNMDYKMDPTQ 368
DB 273 IFAARDTASVMSMLIKLKLAVNPVLEAVTEQNAI-RKDKESGS-LTWDTDKMKPLTS 330

```

```

QY 369 CVINETLRGNVYFLRKLKLDVRYKGYDIPSGMKVLPVISAVALHNSRYDPLENPM 428
DB 331 RVIGETLRVASILSETFREAVEDEVEYGLIPKGMKVLPLFRNIHHSADLFNSNGKRDPS 390
QY 429 RWQOQNNNGASSGSGSFTWGNNTMPFGGPRCLCAGSELAKLEMAVYTHLVKLF 483
DB 391 RFE-----VAPKRNTEPFGNGTHSCPGNELAKLEMSIMIHLLTKY 432

RESULT 14
ID 09F138 PRELIMINARY; PRT; 518 AA.
AC 09F138;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE CYTOCHROME P450-LIKE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=9397451; PubMed=10470850;
RA Kaneo T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
RT Sequence features of the regions of 1,011,550 bp covered by seventeen
RT P1 and TAC clones.";
RL DNA Res. 6:183-195(1999).
DR EMBL: AB017064; BAB1064.1;
SQ
SEQUENCE 518 AA; 58877 MW; CCA3F74F4301563E CRC64;

Query Match 23.5%; Score 631; DB 10; Length 518;
Best Local Similarity 30.5%; Pred. No. 5.7e-37;
Matches 152; Conservative 101; Mismatches 190; Indels 56; Gaps 9;

QY 27 LKRRNKRTPNLPKSGMPFLGETIGYLPYATTLGDPMOOHVSKGYKTVRSNLFGE 86
DB 64 LYRWSNPKCKNKLPRGSMGLPIITGTCDFEPHOLYELSPVKKRMKLYGELFTNTFGS 123
QY 87 PTIYSADAGLNRFILONEGLFECSYRSTIGILGKMSMLVLVGMHRDKSISLNLFLSH 146
DB 124 NTVVLTPEPDIIEFVROENKSEFVSYPDEAFVKGPKGENVFLKHGNIKHVQISLOHGS 183
QY 147 ARLEITILKDVERTTLVLDWMQONSIFSAODEAKKTF-NLAKKHMSMDPGEETEQ 205
DB 184 EALKKMGELDRVTEHLRSKANOGSFDKAEAVESVIMAHLPKIIISNLKPEQAT-L 241
QY 206 KKEVTFMKGVYAPLNPATYHAKALOSRATILKFIERKEER-----LDIKE 255
DB 242 VDNIMALGSEWFOPLMTITLITSTYKFIARVAYLYIKVFTKRKRSREMGDFLDTMV 301
QY 256 EDQEEEVKTEDEAKMSKSDHVRKORTDDLLGVNLKHSNLSTEQIIDLITSLFAGHET 315
DB 302 EEKKEDEVINEES-----AINLIFAILLVAKES 330
QY 316 SVAIALAIFLQCPRAVELREHLEIARAKKELESELNMDYK-KNDFQCYINET 374
DB 331 TTSVTSIAIKFLAENHAKALAEKREHAAIIIONRKGK-AGVSWEYRHOMFTNMTVNET 389
QY 375 LRGNVYRFLRKLKLDVRYKGYDIPSGMKVLPVISAVALHNSRYDPLENPMRMOON 434
DB 390 LRANMAPIMYRKAVNDVEIKGTYTPAGWYAVVPRVHPRDALTAYEPLEPNRMWEGKE 449
QY 435 NGASSGSGSFTWGNNTMPFGGPRCLCAGSELAKLEMAVYTHLVKFMWELEADDPF 494
DB 450 ---LRSGSKTF-----MFGGCVROCVGAEFARLQISIFIHLLVYTDPSLQAESEFI 499

```



OY 495 APEVDFPNCGLPIRVSRII 513  
 Db 500 RAPLFFPKGLPIKISQSL 518

RESULT 15

ID 09LG17 PRELIMINARY: PRT: 735 AA.  
 AC 09LG17  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE F14J16.21.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
 RA Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B., Lee J.,  
 RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC F14J16 from chromosome  
 RT 1."  
 RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RN Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RN Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bel Q., Chin C., Chou J., Choi E., Conn L.,  
 RA Conn L., Conway A., Gonzales A., Hansen N., Howling B., Koo T., Lam B.,  
 RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharzky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -I- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +  
 CC -I- OXIDIZED FLAVOPROTEIN + H(2)O (BY SIMILARITY).  
 CC -I- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM (BY  
 CC SIMILARITY).  
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL: AC002304; AAF79335.1; -  
 DR InterPro: IPR001128; -  
 DR Pfam: PF00067; p450; 2.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 KW Electron transport; Endoplasmic reticulum; Heme; Membrane; Microsome;  
 KW Monooxygenase; Oxidoreductase.  
 SO SEQUENCE 735 AA; 84661 MW; 47186202974F811 CRC64;

Query Match 23.5%; Score 631; DB 10; Length 735;  
 Best Local Similarity 29.8%; Pred. No. 9, 2e-37;

Matches 169; Conservative 99; Mismatches 187; Indels 112; Gaps 15;  
 OY 19 ISLLFLILK-----RRNRKTRPNLPCKSGMPFLGCTGYLKPATATLGPMDQH 71  
 Db 190 LSFVIALVYVVKISLWLYRMANPNCCKLPFGSMGFPVIGETVEFFKPSFNFHPEVKR 249  
 OY 72 VSKY-----GKYSNLFGEPTIVSADAGLNRFITOMEGRLFECGYSRSTIG 118  
 Db 250 MKFLVIFISAGCKNLHGSLFRTNLTGKTIYSTPEVNFETLKQENCFIYSTPEALVR 309  
 OY 119 ILGKSMVLVVG-DMHRDMRSISLNFSLHARLTILKDVNRHTFLVDSMOQNSIFSAQ 177  
 Db 310 IFGKDNLFPGKQDFRVRVRHIALQLGPECKKQRFIQIDIAETSEHLKSVFGQVVDK 369  
 OY 178 DEAKKFTFNLAKKHIMS-MDPEEETEOUKKEYVPMKGVASAP----- 220  
 Db 370 DYSGRILIEOMTIMISNKP--ETRSKILIESFRDFSPDLVRSPPDPSPFNALYNGMLMK 427  
 OY 221 -LNLPGTAVHKALOSRAITLKFIERKMEERKLDIKEEDQEEBEVTEDEAKMSKD 275  
 Db 428 DEVMILNIDPLFY--VKARSVMKMLKMKERREARSDSKYGDPMETIYEYK-- 483  
 OY 276 HVKORTDDDLGWVLRKHSNLSTEOILDLISLPAGHETSSVAIALAIFLQACPKAVE 335  
 Db 484 -----EGDTINEERSVELLSLLASVETSTMTALTVKFAIENPKVLM 527  
 OY 336 ELREHELETA--RAKKELGESLNMDDYK-MDFQCYNIEFLRGVNYRFLHRAKADY 392  
 Db 528 ELKREHETILQNRADKESG--VTWKEYRSMNPFHMTVINESLRGSLSPAFMRANVNDV 584  
 OY 393 RYK-----GYDIPSGMKVLPVISAHLNDSRYDQPN 423  
 Db 585 ELKGRFSLFCYIKITISLPNDLIONRVVAGYITIPAGMIYLVVPSLHYDPQIYEQC 644  
 OY 424 LFNPMRWQOONNAGSSGSGSFTWGNMNYMPFGGPRCLCAGSELAKLENAVFIHLVLF 483  
 Db 645 FENPMWECKE---LISGSKTF-----MARGGARLCAGAEFARLQMAIFLHILVTTY 694  
 OY 484 NMELAEDDQPAFPFVDFPNCGLPIRVS 510  
 Db 695 DESLIDKSYITIRAPLIRFSK--PIRIT 719

Search completed: September 28, 2001, 18:45:40  
 Job time: 130 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:55 : Search time 13.15 seconds  
(without alignments)  
1336.355 Million cell updates/sec

Title: US-09-502-426-2  
Perfect score: 2681  
Sequence: 1 MFETEHHTLPLILPLSLLS.....FAFPVDFPNGLRIVRSRL 513

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues  
Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Swissprot\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	978.5	36.5	472	1	C901_ARATH Q43569 arabidopsis
2	761.5	28.4	464	1	CP85_LYCES Q43147 lycopersico
3	579.5	21.6	490	1	CP83_ARATH Q23051 arabidopsis
4	503	18.8	492	1	CP26_BRARE P79739 brachydanio
5	491.5	18.3	519	1	CP81_MAIZE Q43246 zea mays (m
6	461	17.2	497	1	CP26_MOUSE O55127 mus musculu
7	438	16.3	497	1	CP26_HUMAN Q43174 homo sapien
8	324.5	12.1	520	1	YRV2_CAEBL Q27514 caenorhabdi
9	321.5	12.0	518	1	YRVA_CAEBL Q27514 caenorhabdi
10	316	11.8	500	1	CP12_ARATH Q65790 arabidopsis
11	312.5	11.7	504	1	CP32_RAT P05183 rattus norv
12	312	11.6	501	1	CP15_MOUSE O54749 mus musculu
13	310.5	11.6	520	1	YRV1_CAEBL Q27513 caenorhabdi
14	303	11.3	501	1	CP06_MOUSE Q24750 mus musculu
15	302	11.3	494	1	CPA8_MESAU P24454 mesocricetu
16	299.5	11.2	492	1	CPBC_RAT P33272 rattus norv
17	297	11.1	491	1	CPB9_MOUSE P12790 mus musculu
18	294.5	11.0	491	1	CPB4_RABIT P00178 oryctolagus
19	293	10.9	491	1	CPB1_RAT P00176 rattus norv
20	291.5	10.8	530	1	CP51_YEAST P16614 saccharomyc
21	290.5	10.8	502	1	CP83_ARATH P48421 arabidopsis
22	289	10.8	503	1	CP39_RAT P51538 rattus norv
23	288.5	10.8	492	1	CPB1_MOUSE O55071 mus musculu
24	288.5	10.8	503	1	CP33_HUMAN P05184 homo sapien
25	288.5	10.8	538	1	CP18_DROME Q95078 drosophila
26	288	10.7	459	1	CPXN_ANASP P29980 anabaena sp
27	287.5	10.7	520	1	YRV5_CAEBL Q27510 caenorhabdi
28	287	10.7	506	1	CG92_ARATH Q43602 arabidopsis
29	286.5	10.7	502	1	CP12_HUMAN P51589 homo sapien
30	286.5	10.7	533	1	CP51_CANCA P30859 candida gla
31	284.5	10.6	505	1	C762_SOLME P37122 solanum mel
32	281.5	10.5	504	1	CP3G_MOUSE O64481 mus musculu
33	281	10.5	528	1	CP51_CANTR P14263 candida tro

34	280.5	10.5	504	1	CP3B_MOUSE O64459 mus musculu
35	280	10.4	499	1	C771_SOLME P37123 solanum mel
36	280	10.4	500	1	CP11_RABIT P52786 oryctolagus
37	279.5	10.4	513	1	C773_SOYBN O48928 glycine max
38	279.5	10.4	519	1	YRV8_CAEBL Q27520 caenorhabdi
39	279	10.4	501	1	C4D2_DROME Q27589 drosophila
40	279	10.4	502	1	CPJ3_RAT P51590 rattus norv
41	278.5	10.4	501	1	CP36_RABIT P11707 oryctolagus
42	278	10.4	470	1	CPBK_MOUSE O62397 mus musculu
43	278	10.4	503	1	CP3D_MOUSE O64464 mus musculu
44	277.5	10.4	491	1	CPB5_RABIT P12789 oryctolagus
45	276	10.3	491	1	CPB2_RAT P04167 rattus norv

## ALIGNMENTS

RESULT	ID	Query Match	Score	DB 1	Length
C901_ARATH	Q43569	36.5%	978.5	DB 1	472
AC	Q43569	Best local similarity	41.9%	Pred. No. 1.3e-55	
DT	15-DEC-1998 (Rel. 37, Last sequence update)	Matches	210	Conservative	79; Mismatches 177; Indels 41; Gaps 8;
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	CYTCHROME P450 90A1 (EC 1.14.-.-.)				
GN	CYP90A1 OR CYP90 OR CPD				
OS	Arabidopsis thaliana (Mouse-ear cress)				
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis				
OX	NCBI_Taxid=3702				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. COLUMBIA				
RX	MEDLINE=96200769; PubMed=8612270				
RA	Szekeres M., Nemeth K., Koncz-Kalman Z., Mathur J., Kauschmann A., Altman T., Redei G.P., Nagy F., Schell J., Koncz C.				
RT	"Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in Arabidopsis."				
RL	Cell 85:171-182(1996).				
CC	-I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.				
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
DR	EMBL: X87367; CAA60793.1; -				
DR	EMBL: X87368; CAA60794.1; -				
DR	InterPro: IPR001128; -				
DR	Pfam: PF00067; P450; 2.				
DR	PROSITE: PS00086; CYTOCHROME P450; 1.				
KW	Oxidoreductase; Monooxygenase; Heme.				
FT	BINDING 418 HEME (BY SIMILARITY).				
SQ	SEQUENCE 472 AA; 53785 MW; 41A73F46D4E343F CRC64;				

```

      124 LAKRHSILTMSFANSSTIKHLMIDRLVFNLDSSSVRL--MEAKKTFELTVKQ 181
      192 IMSDPEEEREOLEKKEVFMKGVASAPLNLPGTAHKAQSRATILKFERKEEERKL 251
      182 LMSDPG-EMSESRKEKELVIEGFFSLPLPLESTYTKAIQAR-----KVAEALT 232
      252 DIKEEDDEEVEEKTEDEAEMSKSDHVRKQRTDIDLGVKLKHSNSTEQILDLISLFA 311
      233 VVVKRRREEEGAE-----RKDKMLAALADGDFSEIYDFLALLVA 278
      312 GHESSVAIALAIFLOACPKAVELREHELEIRAKKEGSELNDMDYKKMPTOCVI 371
      279 GYETSTIMTLAVALFETPLALAKLEHEKIRAKMSD--SYSLWSDYKSMPTOCV 336
      372 NETRLGNVVELHRRKALKDVRKGYDIPSGMKVLPVISAHLNDRYDQNLFPWRMQ 431
      337 NETLRANIIIGVRRAMTDEIKGYKIPKMKVFPSSFRVHLDNPHFKARTNPNRMQ 396
      432 QONNGASSSGSFSYMGNN-YMPGGGPRLCAGSELAKLEMAVFIHHLVLKFMELAE 490
      397 -----SNSVYTPSNVFTFPFGGPRLCPEGLARVALSVFLHRLVGFMSVPAEQ 446
      491 DQPAFPVDPENGLPIRVSR 511
      447 DKLVFPPTTOKRYPFVKR 467

      RESULT 2
      CP85_LYCES STANDARD: PRT: 464 AA.
      AC Q43147;
      DT 15-DEC-1998 (Rel. 37, Created)
      DT 15-DEC-1998 (Rel. 37, Last sequence update)
      DE 15-DEC-1998 (Rel. 37, Last annotation update)
      DE CYTOCHROME P450 85 (EC 1.14.-.-) (DWARF PROTEIN).
      GN CYP85 OR D.
      OS Lycopersicon esculentum (Tomato).
      OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
      OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
      OC Solanales; Solanaceae; Solanum.
      OX NCBI_TaxID=4081;
      RN [1]
      RP SEQUENCE FROM N.A.
      RC STRAIN-CV. GCR758;
      RX MEDLINE=96266705; PubMed=8672892;
      RA Bishop G.J., Harrison K., Jones J.J.G.D.;
      RA "The tomato Dwarf gene isolated by heterologous transposon tagging
      RT encodes the first member of a new cytochrome P450 family."
      RL Plant Cell 8:959-969(1996).
      CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
      CC -----
      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
      CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
      CC the European Bioinformatics Institute. There are no restrictions on its
      CC use by non-profit institutions as long as its content is in no way
      CC modified and this statement is not removed. Usage by and for commercial
      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
      CC or send an email to license@isb-sib.ch).
      CC -----
      DR EMBL: U54770; AAB17070.1; -
      DR InterPro: IPR001128; -
      DR Pfam: PF00067; P450.1.
      DR PRINTS: PR00385; P450.
      DR PROSITE: PS00086; CYTOCHROME_P450.1.
      FT Olfactoreductase; Monooxygenase; Membrane; Heme.
      KM BINDING 414 AAE HEME (BY SIMILARITY).
      FT BINDING 414 AAE HEME (BY SIMILARITY).
      SQ SEQUENCE 464 AA; 53706 MW; D2B21AAB7B14B94 CRC64;

```

```

      12 LLLPSLSTLLFLILKRRN--RKTRENLPRGKSGFLGETTGYLKYPTATLG-DFM 68
      5 LIPLSFFGLCIFTALLRMNQVAYKNLPRPGMGMPLRGTEIEFLK-----LGFSFM 58
      69 QOHVSKTKIKYRSNLFGEPITVSADAGINFTLONEGRLEECSPRSIGILGWSMLVL 128
      59 KNQARVGSFEKSHILCCPTIVSDSELNRYILVNEAKGLVPGVQSIDILGKCNIAAY 118
      129 VGDHRRMRSISLNEISHARLRTLLKDVREHNTLFVDSMOQNSIFSADAEKFTENLM 188
      119 NSGAKTKMRGALLSLISPTMRDQLPKIDFMSHLTNW-DNKVYIDQETTKMAFLSS 177
      189 AKHIMSNDPEEET--EOLKKEVYFMKGVASAPLNLPGTAHKAQSRATILKFERK 245
      178 LKQI-----AGIESYSLAQEFSEFFNLVLTSLPINLPTNHYHARGQARKIYNLLRTL 233
      246 MEERKLDKEEDDEEVEEKTEDEAEMSKSDHVRKQRTDIDLGVKLKHS---NLSTEQI 301
      234 IEERR-----ASKETQHDMLGYLMEBEATRFKLTDEM 266
      302 LDLLSLFAGHETSVAIALAIFLOACPKAVELREHELEIRAKKEGSELNDMDY 361
      267 IDLIITILYSGYEVSTSMMAVYKLDHDKVLELKEHNAIEKKRP--EDPIDNDY 324
      325 RSMFTRAVILLETSLRATIYNGVLKRTQDMEINGIILPKGMRIYVYTRELNDYPRLYPD 384
      362 KMDFTQCVINETRLGNVVELHRRKALKDVRKGYDIPSGMKVLPVISAHLNDRYDQ 421
      422 PNLFPWRMQOONNGASSSGSFSYMGNNYMPGGGPRLCAGSELAKLEMAVFIHHLVL 481
      385 PYSFPWRMMDKS-----LEHONSFLVGGGTROCPGKELGVAESTLHLVFT 433
      482 KFMELAEEDQPAFPVDPENGLPIRVSR 510
      434 KYRWEELGGDKLMKFPVEANGLRINVS 462

      RESULT 3
      CP83_ARATH STANDARD: PRT: 490 AA.
      AC Q23051;
      DT 15-DEC-1998 (Rel. 37, Created)
      DT 15-DEC-1998 (Rel. 37, Last sequence update)
      DE 15-DEC-1998 (Rel. 37, Last annotation update)
      DE CYTOCHROME P450 88A3 (EC 1.14.-.-).
      GN CYP88A3 OR YUP8H12.23.
      OS Arabidopsis thaliana (Mouse-ear cress).
      OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
      OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
      OC Brassicales; Brassicaceae; Arabidopsis.
      OX NCBI_TaxID=3702;
      RN [1]
      RP SEQUENCE FROM N.A.
      RC STRAIN-CV. COLUMBIA.
      RA Theologis A., Osborne B.I., Vysotskaya V.S., Federspiel N.A.,
      RA Toriumi M., Yu G., Oji O., Araujo R., Chung E., Dewar K., Dietrich F.,
      RA Ecker J.R., Margalit A., Oelner P., Davis R.W.;
      RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
      RL -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
      CC -----
      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
      CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
      CC the European Bioinformatics Institute. There are no restrictions on its
      CC use by non-profit institutions as long as its content is in no way
      CC modified and this statement is not removed. Usage by and for commercial
      CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
      CC or send an email to license@isb-sib.ch).
      CC -----
      DR EMBL: AC000098; AAB71462.1; -
      DR InterPro: IPR001128; -
      DR Pfam: PF00067; P450.1.
      DR PRINTS: PR00385; P450.

```

Query Match 28.4%; Score 761.5; DB 1; Length 464;  
 Best Local Similarity 33.4%; Pred. No. 9.5e-42;  
 Matches 170; Conservative 90; Mismatches 188; Indels 61; Gaps 10;



0S Zea mays (maize).  
 0C Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:  
 0C Magnoliophyta: Liliopsida, Poales; Poaceae; PACC clade; Panicoideae:  
 0C Andropogoneae: Zea.  
 0X NCBI\_TaxID=4577;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV\_B73;  
 -RX MEDLINE=96004534; PubMed=7549486;  
 RA Winkler R.G., Helentjaris T.;  
 RT "The maize Dwarf3 gene encodes a cytochrome P450-mediated early step  
 in gibberellin biosynthesis.";  
 RL Plant Cell 7:1307-1317(1995).  
 -I- PATHWAY: EARLY STEP IN GIBBERELLIN BIOSYNTHESIS.  
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN ROOTS, DEVELOPING LEAVES, THE  
 CC VEGETATIVE MERISTEM, AND SUSPENSION CULTURE CELLS.  
 CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL: U32579; AAC49067.1; -.  
 DR InterPro: IPR001128; -.  
 DR Pfam: PFO0067; P450; 1.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Transmembrane; Heme.  
 FT TRANSMEM 1  
 FT BINDING 466  
 FT SEQUENCE 519 AA; 57906 MW; 0F8977A024316D95 CRC64;  
 SQ  
 Query Match 18.3%; Score 491.5; DB 1; Length 519;  
 Best Local Similarity 28.2%; Pred. No. 2e-24; Indels 49; Gaps 11;  
 Matches 129; Conservative 91; Mismatches 189;  
 QY 36 RNLPPGKSGWPLFETIGYLNKPYATATLGDGMOOHVSKYGR-IYRSNLFGEPIT-USA 92  
 Db 69 RARLPPEGMGWPLVGGMMAFLRAFKSGKRDARLASVRRFGTGYRSMFSSPIYLVTT 128  
 QY 93 DAGLNFLIÖNGRLEECYPRISIGSICGLKWSMLVLVGMHDMKMSISLNFSHARLRTI 152  
 Db 129 AECCCKLVMDDDA--FVTMPKRAYALVLRPSFVAMPYDEHRRIRKLTAPINFDALDG 186  
 QY 153 LKLDVRRHRLFLVDSW-QQNSIFSAODEAKKFTFNLMAHINSMDBGEETQLKKEYT 211  
 Db 187 YLRFIDRYTSSIRAMADHGGVVEFLTELRKMTFKIYV-IFLGADQATTLALERSYE 245  
 QY 212 FMKGVASRLNPGTAUHNKALÖSRATILKFERKKERKLDIKE-----EDÖEEVEYKT 265  
 Db 246 LNYGMRAAMINLPFGAYRGALRARRLVAVLÖGVDERBARAKVSGGGVMDRLLEA 305  
 QY 266 EDÖAEKSKDHRKQKOTDDDLGCVLKHNSLSTEOILLDLISLFRAGHTSSVALALAI 325  
 Db 306 QDE-----RGRLLDD-----ETIDLVMYLNAHGSSGHITMMATV 342  
 QY 326 FLÖACPKAAVEELREELLETARAKKELGESELNMDWDYKKMDFOCYINELRLGNVRFLL 385  
 Db 343 FLÖÖNDMEARAKKAEÖELMRSTPS-SQSGILTLRFQRKKMEYLSÖYIDETLRILVNSPVSF 401  
 QY 386 RALKDVRYYKYDIPSGKVLVPLVISAHLNDSRYDÖPLNFWMRWOÖÖNNGASSGSGSF 445  
 Db 402 RQATRDVFNVGYSILPRGKMYQLMYSVNDHPÖVYRDPRTKFDESWE-----GHSPRAGTF 456  
 QY 446 STYMGNNYMPFGGPRICASELAKLEMAFTNHLVLFK 483  
 Db 457 -----LAFGLGARLCSPNDLAKLETISVLEHNHFLFGY 487

```

RESULT 6
CP26_MOUSE STANDARD: PRT: 497 AA.
ID CP26_MOUSE 053127
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYTOCHROME P450 26 (EC 1.14.-.-) (RETINOIC ACID-METABOLIZING
DE CYTOCHROME (P450RA1) (RETINOIC ACID 4-HYDROXYLASE).
GN CYP26A1 OR CYP26 OR P450RA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=C3H;
RX MEDLINE=97392446; PubMed=9250660;
RA Fujii H., Sato T., Kaneko S., Gotoh O., Fujii-Kuriyama Y., Osawa K.,
RA Kato S., Hamada H.;
RT "Metabolic inactivation of retinoic acid by a novel P450
RT differentially expressed in developing mouse embryos."
RL EMO J. 16:4163-4173(1997).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=98113212; PubMed=9442090;
RX Abu-Abed S.S., Beckett B.R., Chiba H., Chithalen J.V., Jones G.,
RA Metzger D., Chambon P., Petkovich M.;
RT "Mouse P450RA1 (CYP26) expression and retinoic acid-inducible retinoic
RT acid metabolism in F9 cells are regulated by retinoic acid receptor
RL gamma and retinoid X receptor alpha."
RL J. Biol. Chem. 273:2409-2415(1998).
CC -1- FUNCTION: PLAYS A KEY ROLE IN RETINOIC ACID METABOLISM. ACTS ON
CC RETINOIDS, INCLUDING ALL-TRANS-RETINOIC ACID (RA) AND ITS
CC STEREOISOMER 9-CIS-RA. CAPABLE OF BOTH 4-HYDROXYLATION AND 18-
CC HYDROXYLATION. RESPONSIBLE FOR GENERATION OF SEVERAL HYDROXYLATED
CC FORMS OF RA, INCLUDING 4-OH-RA, 4-OXO-RA, AND 18-OH-RA.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC -1- INDUCTION: BY RETINOIC ACIDS (RA).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; Y12657; CA73206.1; -.
CC MCD; MGI:1096359; CYP26.
CC InterPro: IPR001128; -.
CC Pfam: PF000385; P450. 1.
CC PRINTS: PR00385; P450.
CC PROSITE: PS00086; CYTOCHROME_P450. 1.
CC Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
CC Endoplasmic reticulum.
CC BINDING 442 442 HEME (POTENTIAL).
CC SEQUENCE 497 AA; 56177 MW; 33B07D7C29134471 CRC64;
Query Match 17.2%; Score 461; DB 1; Length 497;
Best Local Similarity 28.4%; Pred. No. 1.6e-22;
Matches 148; Conservative 78; Mismatches 231; Indels 64; Gaps 14;
OY 10 LPELLPLSL-----LSLLFLILLK-----RRNRKTRFNLPPCKSGMPFGETITGYLK 57
Db 3 LPALASALCTGFVLPFLFLFLAALKLMDLYCVSSRRORSALPLPCTMGPFPGGETL----- 58
OY 58 PYTATTLDDPMQOHNSKGTGIYRSNLFGEPITVASDAGLNRIITLNEGLFPCSYPRSIG 117
Db 59 -QMVLRKRRFLQMKRRKRGFTYKTLFEGPPYRVAGADNVRILLGERRLVSVHPAPSVR 117

```

QY 118 GILKWSMLVGDHMRMSISINFLSHARLRTLLKDVREHRTFLVDSMOQ--NSIFS 175  
 DB 118 TILAGCSNLDHSSHKORRKYIMQAFREALQCYLV-IAEVSQCEQLSGCERLL 176  
 QY 176 AODEPKEFLMAKHSMDPE---EETEDLKEEYVTEPMKGVASAPLPGTAHYKA 231  
 DB 177 VYPEKRLMFIAIRILLGCEPAGGEDQOLVAFEEEMTRNLFSPIDVPSGLYRG 236  
 QY 232 LQSRATILKFERKME--RKLDIKEED--QEEEVKTEDEAEKSDHNRKQRTDDL 286  
 DB 237 VKARNLHARLEENIRAKIRLQATEPBGCKDLQLLIEHSWE-----RGERLDMQ- 288  
 QY 287 LGWVLKHSNLTEDILDLILSLFAGHETSSVATAIAIFFOACPKAVEELREHLEIAR 346  
 DB 289 ---ALKQS--STE-----LIFGHEHTTASATSLITYGLVPHVYLQKREELKSKGL 335  
 QY 347 AKKELGESELMWDYKKMDPTQCVINETLRIGNVVRLHRRALDVRKGYDIPSGMKVL 406  
 DB 336 LCKSNODKRLMETLEQLKTYIGCVAKETRLNRPVPGGFVALKTFEELNGYQIPKGMVVI 395  
 QY 407 PVISAVHLDNSRYDQPNLFNPMWMOQNNGASSSGSFSFTWGNMYMPFGGPRCLASE 466  
 DB 396 YSIDTDFHVDADIFTNKEEFNDRFLVPHEDASRFS-----FIFPGGLRSCVGEKE 446  
 QY 467 LAKLEMAVFIHHLVLFKFMWELAE-----DQPFAPFPVDFP 502  
 DB 447 FAKILKLFYVELARHCWQMLNCPPTMKTSPTVYVNDLP 487

RESULT 7  
 CP26\_HUMAN  
 ID CP26\_HUMAN STANDARD: PRT: 497 AA.  
 AC 043174:

DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CYTOCHROME P450 26 (EC 1.14.-.-) (RETINOIC ACID-METABOLIZING  
 DE CYTOCHROME) (P450RA1) (RETINOIC ACID 4-HYDROXYLASE).  
 GN CYP26A1 OR CYP26  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-97373542; PubMed-9228017;  
 RA White J.A., Beckett-Jones B., Guo Y.-D., Dilworth F.J., Bonasoro J.,  
 RA Jones G., Pelkovich M.;  
 RT "cDNA cloning of human retinoic acid-metabolizing enzyme (hP450RA1)  
 RT identifies a novel family of cytochromes P450.";  
 RL J. Biol. Chem. 272:18538-18541(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-98380037; PubMed-9716180;  
 RA Someveld E., van den Brink C.E., van der Leede B.M., Schultes R.K.,  
 RA Petkovitch M., van der Burg B., van der Saag P.T.;  
 RT "Human retinoic acid (RA) 4-hydroxylase (CYP26) is highly specific for  
 RT all-trans-RA and can be induced through RA receptors in human breast  
 RT and colon carcinoma cells.";  
 RL Cell Growth Differ. 9:629-637(1998).  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RA MEDLINE-99045433; PubMed-9826557;  
 RA Trofinova-Griffin M.E., Juchau M.R.;  
 RT "Expression of cytochrome P450RA1 (CYP26) in human fetal hepatic and  
 RT cephalic tissues.";  
 RL Biochem. Biophys. Res. Commun. 252:487-491(1998).  
 CC -1- FUNCTION: PLAYS A KEY ROLE IN RETINOIC ACID METABOLISM. ACTS ON  
 CC RETINOIDS, INCLUDING ALL-TRANS-RETINOIC ACID (RA) AND ITS  
 CC STEREOISOMER 9-CIS-RA. CAPABLE OF BOTH 4-HYDROXYLATION AND 18-  
 CC HYDROXYLATION. RESPONSIBLE FOR GENERATION OF SEVERAL HYDROXYLATED  
 CC FORMS OF RA, INCLUDING 4-OH-RA, 4-OXO-RA, AND 18-OH-RA.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.

CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN ADULT LIVER, HEART,  
 CC PITUITARY GLAND, ADRENAL GLAND, PLACENTA AND REGIONS OF THE BRAIN.  
 CC -1- INDUCTION: BY RETINOIC ACIDS (RA).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF005418; AAB88881.1; -  
 CC DR MIR: 602239; -  
 CC DR InterPro: IPR001128; -  
 CC DR Pfam: PF00067; P450; 1.  
 CC DR PRINTS: PR00385; P450.  
 CC DR PROSITE: PS00086; CYTOCHROME P450; 1.  
 CC KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;  
 CC Endoplasmic reticulum.  
 CC FT BINDING 442 442 HEME (POTENTIAL).  
 CC SEQUENCE 497 AA; 56162 MW; EAB6B84B24B2EAB3 CRC64;

Query Match 16.3%; Score 438; DB 1; Length 497;  
 Best local Similarity 27.0%; Pred. No. 4.9e-21; Mismatches 226; Indels 72; Gaps 15;  
 Matches 142; Conservative 85;

QY 10 LPLLLPSL---LSLLFLILLK-----RRNRKTRFNLPPGSGMPFLGRTGYLK 57  
 DB 3 LPALLASALCTFVPLDLFLFLAIAIKMDLYCSGRSCALPLPCTMGPFFRGELT----- 58

QY 58 PYTATTLDEPMQOHVSKGKTYRSMLECEPTIVSADAGLNRFITLONERLECESTPRSIG 117  
 DB 59 -QWVLORRKFQMKRRKCYFYKTHLFGPRVYVAGANVRITLLGDRLVSVHWPASVR 117

QY 118 GILKWSMLVGDHMRMSISINFLSHARL--TLLDVERHRTFLVDSMOQ--NS 172  
 DB 118 TILAGCSNLDHSSHKORRKYIMQAFREALQCYLV-IAEVSQCEQLSGCERLL 176

QY 173 IFSADENAKKETFLMAKHSMDP-----GEETEDLKEEYVTEPMKGVASAPLPGTA 227  
 DB 174 GLVYPEVKRLMFIAIRILLGCEPAGGEDQOLVAFEEEMTRNLFSPIDVPSG 232

QY 228 YHKAQSRATILKFERKMEERKLDIKEED-----QEEEVKTEDEAEKSDHNRKQRT 282  
 DB 233 LYRGKAKARNLHARLEENIRAKIRLQATEPBGCKDLQLLIEHSWE-----RGERL 285

QY 283 DDDLGLWVLKHSNLTEDILDLILSLFAGHETSSVATAIAIFFOACPKAVEELREHLEIAR 342  
 DB 286 DMQ-----ALKQS--STE-----LIFGHEHTTASATSLITYGLVPHVYLQKREELK 331

QY 343 ETARAKKEGESELMWDYKKMDPTQCVINETLRIGNVVRLHRRALDVRKGYDIPSG 402  
 DB 336 LCKSNODKRLMETLEQLKTYIGCVAKETRLNRPVPGGFVALKTFEELNGYQIPKGMVVI 395

QY 403 WKVLPVISAHLNDSRYDQPNLFNPMWMOQNNGASSSGSFSFTWGNMYMPFGGPRCLASE 462  
 DB 392 MWVYISIDTDFHVDADIFTNKEEFNDRFLVPHEDASRFS-----FIFPGGLRSC 442

QY 463 AGSEILAKLEMAVFIHHLVLFKFMWELAE-----DQPFAPFPVDFP 502  
 DB 443 VKEPERAKILKLFYVELARHCWQMLNCPPTMKTSPTVYVNDLP 487

RESULT 8  
 YRV2\_CAEEL  
 ID YRV2\_CAEEL STANDARD: PRT: 520 AA.  
 AC 027514:  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)





QY 289 WLK-HSNLSTEOILDLILSLFAGHETSVAIALAIFLOACPKAVEELREHLEIARA 347  
 DB 303 EVLAKVDKLTDELIGOLFVFLAGYDTTALSSSYLLATLHPKIOKLEE-----V 356  
 QY 348 KKLGESELMNDYKKMDFCTOCVINEITRLGNVVRFLH-RALKDVRKYGDIPDSGMKVL 406  
 DB 357 DRECPDEPVEFTDOLSKLTYECVKEALRYPLASLVHNRKCLTKTNVIGMEIEAGTNIN 416  
 QY 407 PVISAVHLNDRY-DQPLNLFPMRQOONNAGSSGSFSTWGNVMPFGGPRCLAGS 465  
 DB 417 VDTSLHDPKVPWGDVNEFKPERE-----SGDELFPKAG-GYLPFGKPRICIGM 467  
 QY 466 ELAKLEMAVFIHLVLKFNWE 486  
 DB 468 RLAMMEMKMLTJNLKNTFE 488

RESULT 10  
 C912 ARATH STANDARD; PRT: 500 AA.  
 AC 065790: 09S2U3:  
 DF 01-OCT-2000 (Rel. 40, Created)  
 DF 01-OCT-2000 (Rel. 40, Last sequence update)  
 DF 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE CYTOCHROME P450 91A2 (EC 1.14.-.-).  
 GN CYP91A2 OR AT4G37430 OR F6G17.80.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV\_COLUMBIA; TISSUE=Seedling;  
 RX MEDLINE=96281573; PubMed=9620263;  
 RA Mitutani M., Ward E., Ohta D.,  
 RT "Cytochrome P450 superfamily in Arabidopsis thaliana: isolation of  
 RT cDNAs, differential expression, and RFLP mapping of multiple  
 RT cytochromes P450."  
 RL Plant Mol. Biol. 37:39-52(1998).  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV\_COLUMBIA;  
 RX MEDLINE=20083488; PubMed=10617198;  
 RA Mayer K.F.X., Schueller C., Mambutt R., Murphy G., Volckaert G.,  
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,  
 RA Harris B., Ansoyge W., Brandt P., Grivell L., Rieger M.,  
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,  
 RA Kreis M., Delseny M., Puldomenech P., Watson M., Schmidheini T.,  
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
 RA Vos P., Hohenseil J., Zimmermann W., Medler H., Ridley P.,  
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,  
 RA Van der Schueren J., Glymonprez B., Chuang T.-C., Vandembussche F.,  
 RA Breken M., Welfens I., Voet M., Bastiaens I., Aert R., Dettor E.,  
 RA Wetzelsieger T., Bothe G., Ramsperger U., Hilbert H., Bruun M.,  
 RA Holzer E., Brandt A., Peters S., van Stevenen M., Dikse W.,  
 RA Meulman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,  
 RA Bernieser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
 RA De Keyser A., Buysnaert C., Giejen J., Villarroel R., De Clercq R.,  
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,  
 RA Pettett A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,  
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehnert T.-H.,  
 RA Dose S., de Haan M., Maarse A., Schaefer M., Meller Auer S.,  
 RA Gabel C., Fuchs M., Fattmann B., Granderath K., Dauner D., Herzl A.,  
 RA Neumann S., Argitidou A., Vitale D., Liguori R., Piravandi E.,  
 RA Massenat O., Quidley F., Clabaud G., Muendlein A., Felber R.,  
 RA Schnabel S., Hillier R., Schmidt W., Lecharny A., Aubourg S.,  
 RA Chedid F., Cooke R., Berger C., Monfort A., Casacuberta E.,  
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,  
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,  
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,

RA Zaccaria P., Beyan M., Wilson R.K., de la Bastide M., Habermann K.,  
 RA Parnell L., Dethia N., Gnoj L., Schütz K., Huang E., Spiegel L.,  
 RA Sekon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,  
 RA Stoneking T., Kalicki J., Graves T., Hamon G., Edwards J.,  
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K.,  
 RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
 RA Krimer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,  
 RA Nelson J., Spith J., Ryan E., Andrews S., Giesel C., Layman D.,  
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,  
 RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,  
 RA Ma P., Zhong J., Preston R., Vil D., Shekhar M., Matero A., Shah R.,  
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lodhi M., Johnson A.,  
 RA Chen E., Marra M., Martienssen R., McCombie W.R.,  
 RA "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
 RA thaliana."  
 RL Nature 402:769-777(1999).  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: D76607: BAA28539.1; -  
 DR EMBL: AL035601: CAB38210.1; -  
 DR EMBL: AL015591: CAB80408.1; -  
 DR Mendel: 29894; Arath:1113.29894.  
 DR InterPro: IPR001128; -  
 DR Pfam: PF00067: P450.1;  
 DR PROSITE: PS00086; CYTOCHROME\_P450.1;  
 KW Oxidoreductase; Monooxygenase; Membrane; Heme.  
 FT BINDING 438  
 FT CONFICT 106 A->V (IN REF. 1).  
 FT CONFICT 127 N->I (IN REF. 1).  
 FT CONFICT 140 N->I (IN REF. 1).  
 FT CONFICT 454 454. S->T (IN REF. 1).  
 SQ SEQUENCE 500 AA: 57555 MW: 6745352070EA2EA CRC64:

Query Match 11.8%; Score 316; DB 1; Length 500;  
 Best Local Similarity 25.2%; Pred. No. 3.1e-13;  
 Matches 135; Conservative 83; Mismatches 199; Indels 118; Gaps 22;

QY 9 LPLPLLSLSLFLILKLR--NRKTRFNLPPKSGWPLGETIGYIKRYATATLGD 66  
 DB 1 MLYFILP-----LFLVSYKFEKSKQRFNLPGPSPRPVGH-LHLMKP-----PIHR 50  
 QY 67 FMOQVSKYKGRYRNLNGEPTIV-----SADAGLNRIILONEGRFLFCSPRSIG 117  
 DB 51 LLQRSNQYGPILF-SLRFSRRVYVITSPSLAOSEFTGNDVLSSRPLQTLAKVAVYMH 109  
 QY 118 GLIGKSMVLVGDHNRDRI-SLNFSLHARLRT--LTKDVERHTFLVDSMOONSIF 174  
 DB 110 TTIVG-----TAPGDMRWNRILRISQELISSHRLINPQIRKDEILMLTRLRLRYQTS-- 163  
 QY 175 SAQDEAKRFT-----NLAKHIMSMPG-----EETEOIKREYVFMKVQ 216  
 DB 164 ---NSNDFTHLEPLSLDPLFNNIVRWVTKRYGGDVNKKKEAELEPKK----- 211  
 QY 217 VSAPLNLPGTAVHKLQSPAT-----LKTIERKMERKLDIYEDDEEVEVTEDEA 269  
 DB 212 -----LVYDIAMYSQANHSADYLPILTKLFGKFEKVEVAIG-----KSMDDI 253  
 QY 270 ESKSGDHVKOKTDDDLGLGVKHSNLSLEQILD-----LISLFPAGHETSSVAIALAI 324  
 DB 254 LRLLDLDECRDKRKGNTVNHVLSLDQOOPRYTVDYIKLMSMMLAGETSAVLENNAA 313  
 QY 325 FLQACPKAVEELREHLEIARAKKEGSEL--NMDDYKKMDFCTOCVINEITRLGNVVR 383

Db 314 ANLRNEVELEKARSE-----IDKIGKORLIDESDIAVLPIQNVSEFRLFPVAPF 367  
QY 384 L-HRKALKDYRYKGYDIPSGMKVLPVISAHLNDSRYDOPNLFNPMWQOONNGASSGS 442  
Db 368 LIPSPFTDMKIGGYDVPRIYVWNAIAHRLDEIWEPEKFNPDYR---NDGC----- 419  
QY 443 GSPFTWNNY-----MFGGGRPLCAGSELAKLEMAVFIHHLVLFKFWELAEDDQ 492  
Db 420 -----GSDYVYVKLMPFGNGRRTCPGAGORIVTLATLSLIOCFFEMWVKGE 468  
RESULT 11  
CP32\_RAT STANDARD: PRT: 504 AA.  
ID CP32\_RAT P05183; Q64672; Q64629;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DE CYTOCHROME P450 3A2 (PC 1.14.14.1) (CYP11A2) (P450-PCN2) (P450/6-  
DE BETA-A) (TESTOSTERONE 6-BETA-HYDROXYLASE).  
GN CYP3A2 OR CYP3A-2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87064606; PubMed=3785219;  
RA Gonzalez F.J., Song B.-J., Hardwick J.P.;  
RT "Pregnenolone 16 alpha-carbonitrile-inducible P-450 gene family: gene  
RL conversion and differential regulation";  
RL Mol. Cell. Biol. 6:2969-2976(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
RC MEDLINE=91254339; PubMed=2043144;  
RA Miyata M., Nagata K., Yamazoe Y., Kato R.;  
RT "A gene structure of testosterone 6 beta-hydroxylase (P45011A).";  
RL Biochem. Biophys. Res. Commun. 177:68-73(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
RC MEDLINE=95070139; PubMed=7979376;  
RA Miyata M., Nagata K., Shimada M., Yamazoe Y., Kato R.;  
RT "Structure of a gene and cDNA of a major constitutive form of  
RL testosterone 6 beta-hydroxylase (P450/6 beta A) encoding CYP3A2:  
RL Arch. Biochem. Biophys. 314:351-359(1994).  
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY  
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
CC ACIDS, AND XENOBIOTICS.  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, ENDOPLASMIC RETICULUM.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: M13646; AAA41051.1; -  
CC EMBL: X79319; CAAS5887.1; -  
CC EMBL: X79320; CAAS5888.1; -  
CC EMBL: U09742; AAAB2168.1; -  
CC EMBL: U09734; AAB60492.1; -  
CC EMBL: U09725; AAB60492.1; JOINED.

DR EMBL: U09726; AAB60492.1; JOINED.  
DR EMBL: U09727; AAB60492.1; JOINED.  
DR EMBL: U09728; AAB60492.1; JOINED.  
DR EMBL: U09729; AAB60492.1; JOINED.  
DR EMBL: U09730; AAB60492.1; JOINED.  
DR EMBL: U09731; AAB60492.1; JOINED.  
DR EMBL: U09732; AAB60492.1; JOINED.  
DR EMBL: U09733; AAB60492.1; JOINED.  
DR PIR: A25222; A25222.  
DR HSSP: P14779; 1FAG.  
DR InterPro: IPR001128; -  
DR InterPro: IPR002397; -  
DR InterPro: IPR002401; -  
DR InterPro: IPR002402; -  
DR InterPro: IPR002403; -  
DR Pfam: PF00067; P450. 1.  
DR PRINTS: PR00385; BP450.  
DR PRINTS: PR00385; P450.  
DR PRINTS: PR00463; EP450I.  
DR PRINTS: PR00464; EP450II.  
DR PRINTS: PR00465; EP450IV.  
DR POSITE: PS00086; CYTOCHROME P450; 1.  
KW oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 443 443 HEME.  
FT CONFLICT 56 57 GL -> AV (IN CAAS5887).  
FT CONFLICT 429 429 D -> H (IN REF. 1).  
FT CONFLICT 445 445 G -> D (IN REF. 1).  
SQ SEQUENCE 504 AA; 57731 MW; C2C84AF736035AB2 CRC64;

Query Match 11.7%; Score 312.5; DB 1; Length 504;  
Best local similarity 25.5%; Pred. No. 5, 2e-13;  
Matches 137; Conservative 96; Mismatches 211; Indels 93; Gaps 24;

QY 9 LLELLLPSTLSLLPLILK--RNKRTFNP--PGKGMFLETGTCTYKLRATFTL 64  
Db 3 LLSALTEWLVLLVLLVLLRGLGTHRGIFKKOGIPGKPLFGLVLYR----GL 57  
QY 65 GDFMOQHVSKYKIRYSNLSNLEPTIVSA--DAGLRFILQNEGRLEEC---SYPSISG 117  
Db 58 GRDMDEKYYKKYKIV--GLFDGQTPVPAIMDEMIKKVL-----VACGFVFNRRRFGP 110  
QY 118 -GLLKSMVLVLDGMHRMRSISLNFSLHARLTILKQVRRHTLVLDVSDQONS---- 172  
Db 111 VGIMKRAVSAKDEEMKRYRALLSPFTS--GRUKE--MFLIEYGDILVYKLOAEETGR 168  
QY 173 -----IFSA--ODEAKKFTFNLMAKHMS--MDPGEETEOLK--EYVFMKGVYAPL 221  
Db 169 PVIMKVVFGYSMDVITTSFGVNVDSLNPKDPFEVETIKLLRFDFDFLFSVLFPP 228  
QY 222 NLEGTAVHKA-----OSRATILKFERKMEERKLIDNEEDQEEVEVTEDEAEMSKS 274  
Db 229 LTP---IYELNLCMPRKOSIAFFQFVHR--IKETRLDSNKHRR-----VDFLDMLN 277  
QY 275 DHRKQRTDDDLGLWLNKSLSTBOILLILSLFRAGHTSSVAALALAFILOAPKAV 334  
Db 278 AH---NNSKDE---VSHNALSVDLIAOSVIFRAGYVTTSTTSFVLYFATIHDIQ 329  
QY 335 EELEREHELEIARAKKEGESELMWDDYKKMDFTQCYINETFLRGVNYFLHRAKLDVRY 394  
Db 330 KRIQEE-IDCALPCK---APPTDYIEMENYLDWLNELRLRPLIGNRLERCKADIEL 384  
QY 395 KGYDIFSGMKVLPVISAHLNDSRYDOPNLFNPMWQOONNGASSGSFSTGNNYMP 454  
Db 385 DGLFIRKGSVVTIPYALHHDHPQWPKRPEFHRERSKEKKSIDP-----VYILP 435  
QY 455 FGGGPRPLCAGSELAKLEMAVFIHHLVLFKFWELAEDDQPPAFVDFPNCPLTRVSR 511  
Db 436 FGGGPRNCIMRPLALNMKLTALTKVLQNSFQPKETQ-----IPLKSR 480

RESULT 12

CPJ5\_MOUSE  
ID CPJ5\_MOUSE STANDARD; PRT; 501 AA.  
AC 054749;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE CYTOCHROME P450 2J5 (EC 1.14.14.1) (CYPIJ5) (ARACHIDONIC ACID  
DEPOXYGENASE).  
GN CYP2J5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;  
RX MEDLINE=98234557; PubMed=9570962;  
RA Ma J., Ramachandran S., Fiedorek F.T. Jr., Zeldin D.C.;  
RT Mapping of the CYP2J cytochrome P450 genes to human chromosome 1 and  
mouse chromosome 4.";  
RL Genomics 49:152-155(1998).  
CC -1 CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC -1 OXIDIZED FLAVOPROTEIN + H(2O).  
CC -1 SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U62294; AAB87635.1; -  
DR MGD; MG1:1270149; CYP2J5.  
DR InterPro: IPR001128; -  
DR InterPro: IPR002401; -  
DR Pfam: PF00067; P450.1.  
DR PRINTS: PR00385; P450.  
DR PRINTS: PR00463; EP4501.  
DR PROSITE: PS00086; CYTOCHROME\_P450.1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Mitochondrion; Endoplasmic reticulum.  
FT BINDING 447 447 HEME (BY SIMILARITY).  
FT SEQUENCE 501 AA; 57784 MW; C67F2E79DD64AF99 CRC64;  
SQ  
Query Match 11.6%; Score 312; DB 1; Length 501;  
Best Local Similarity 23.0%; Pred. No. 5.5e-13;  
Matches 117; Conservative 96; Mismatches 203; Indels 92; Gaps 17;  
OY 14 LRLPILSLFLILIKRRNRKTRFNLPRGSGWPLGFTIGYLKPYATTLGDMQOHVS 73  
D 1 LVLVAVVFLFLINILIRSHNRK---NYPGPRWLRPFVGNFQIDTKQTHLVL---QGVK 73  
OY 74 KYGIVRSNLEGEPTIYSADAGL-----NFIIONERLRECEYPSRISGIL 120  
D 74 KYGVVFSLDLQSPVYVSGPLIKEMETHLDONFVNRFKPRVR-----IT 121  
OY 121 GKWSMLVLVGMHRDMRSISINFLSHARL-RTILKDYERTTLFVLSMQONSIFSAD 179  
D 122 GKNGLVVSNGOTWKEORLALMRNGLGKKSLEERTQETNHLVLAIREGGOPRPH 181  
OY 180 AKKTFNLMAKHMSMDPE-----EETEDLKKEYVTFMGVSAPLN----- 222  
D 182 LK--LINNVSNILISVYGERFEDYEDCOFELLQILDETMLMGSSAGOLYNGFPCIMKY 239  
OY 223 LPTGAVHALGSRATILKIERKMEERKLDITKEEQEEEEKTEDEA---EMSKSDHVRK 279  
D 240 LFG--PHQKIFRNMGKLIKFSHVYKHK-----EKDMNPDEPRDIDAVLLEMK----- 287  
OY 280 QRTDDLLGWLKHSNLSSTEOILLDLISLFLAGHETSSVAIALAIFFLQACPKAVEILRE 339

DB 288 ---DPD-----RTSFMENISITLLEFLGCTETTSITLRLMALLVNSVPEIOENQA 338  
OY 340 EHLFARAKKKEIGSELNMDYKKMDFQCYINTEFLRGVNR--LHRKALKDVRKYKD 398  
D 339 EIDRVIGHKROYLSL-----DRESMPYTNNAVTHEOVRGNITVPLNLSREVVDTKFNHF 393  
OY 399 IPSGKVLPLVSAVHLDSRSDQPLFPMWRHQOONNASSSGSGSFTWGNVMPFGG 458  
D 394 LPKGMILTINLALRDRKEMATPEVFNHEFL-----NQGFK-RESFLPSMG 443  
OY 459 PRLCAGSELAKLEMAVFIHLHLVKNWE 486  
D 444 KRACIGEQALAKSELFIFPSALMKRTFR 471  
RESULT 13  
YRVL\_CAEEL  
ID YRVL\_CAEEL STANDARD; PRT; 520 AA.  
AC 027513;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PUTATIVE CYTOCHROME P450 CYP13A4 (EC 1.14.-.-).  
GN CYP13A4 OR T10B9.1.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Gardner A.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1 FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOATE  
CC MONOOXYGENASES. THEY OXIDIZE A VARIETY OF STRUCTURALLY UNRELATED  
CC COMPOUNDS, INCLUDING STEROIDS, FATTY ACIDS, AND XENOBIOTICS.  
CC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z48717; CAAB8603.1; -  
DR WormPep: T10B9.1; CE01654.  
DR InterPro: IPR001128; -  
DR InterPro: IPR002402; -  
DR Pfam: PF00067; P450.1.  
DR PRINTS: PR00385; P450.  
DR PRINTS: PR00464; EP45011.  
DR PROSITE: PS00086; CYTOCHROME\_P450.1.  
KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme.  
FT BINDING 464 464 HEME (BY SIMILARITY).  
FT SEQUENCE 520 AA; 59371 MW; 996C1172B8D64E93 CRC64;  
SQ  
Query Match 11.6%; Score 310.5; DB 1; Length 520;  
Best Local Similarity 23.6%; Pred. No. 7.2e-13;  
Matches 121; Conservative 102; Mismatches 177; Indels 113; Gaps 23;  
OY 42 GKSGWPLG-----ETIGYLKPYATTLGDMQOHVSRYKGYRSNLEGEPTIYSAD 93  
D 33 GPRGPFPGVILIKFHDYDNPGLK-----LGEM-----TKYGSYIGITEGVKTLVSN 82  
OY 94 AGLNFIIONGRLEFECYPSISGILG-----KSMVLVY-GDMHRDMRSISINFLSHA 147  
D 83 P---EFVHEVVKQFDNRYGKRTNPIQDPPKKNRAHLVLAQGHMKRLKRTLASPTFSKH 139  
OY 148 RLRTIL-----LKVVERH-----TLFVLSMQONSIFSADAKKTFNLMAKHI 192







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 : Search time 19.47 Seconds  
(without alignments)  
2007.064 Million cell updates/sec

Title: US-09-502-426-2

Sequence: 1 MFETEHNTLLPLLPILLS.....FAFPVDPNGLPIRVSRLL 513

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2677	99.9	513	2	T46143	steroid 22-alpha-h
2	978.5	36.5	472	1	S55379	cytochrome P450 cy
3	866	32.3	512	2	H96759	probable steroid 2
4	860	32.1	457	2	D85429	cytochrome P450 1i
5	815	30.4	382	2	T48613	hypothetical prote
6	761.5	28.4	464	2	T07859	cytochrome P450 ho
7	632	23.6	457	2	T04444	cytochrome P450 -
8	611	22.8	482	2	T02739	probable cytochrom
9	581	21.7	487	1	C71417	cytochrome P450 d1
10	579.5	21.6	485	2	A84859	probable cytochrom
11	579.5	21.6	490	2	H86185	hypothetical prote
12	577	21.5	455	2	T48973	cytochrome P450-1i
13	547.5	20.4	489	2	D96813	probable cytochrom
14	497.5	18.6	460	2	D96813	hypothetical prote
15	491.5	18.3	519	1	T02263	cytochrome P450 DW
16	489.5	17.9	444	1	S75761	cytochrome P450 ho
17	479.5	17.3	444	1	S75761	hypothetical prote
18	369	13.8	349	2	A86329	hypothetical prote
19	357	13.3	518	2	T20908	hypothetical prote
20	338	12.6	453	2	C83722	cytochrome P450 hy
21	329.5	12.3	517	2	T20907	hypothetical prote
22	324.5	12.1	520	2	T24778	hypothetical prote
23	321.5	12.0	518	2	T24783	hypothetical prote
24	316	11.8	500	2	T04737	cytochrome P450 ho
25	310.5	11.6	520	2	T24777	hypothetical prote
26	310	11.6	500	2	T52175	cytochrome P450 mo
27	306.5	11.4	504	2	A25222	cytochrome P450 3A
28	302.5	11.3	491	2	S31277	cytochrome P450 2B
29	300	11.2	491	2	T84735	testosterone 16a-h

30	299.5	11.2	492	2	S27160	cytochrome P450 2B
31	297	11.1	491	2	A31047	testosterone 16alp
32	297	11.1	494	2	A33293	cytochrome P450 2A
33	294.5	11.0	491	1	O4RBP	cytochrome P450 2B
34	294	11.0	506	2	D96672	probable Cytochrom
35	294	11.0	516	2	T48140	flavonoid 3',5'-hy
36	292	10.9	491	1	O4RBP	cytochrome P450 2B
37	291.5	10.9	491	1	S35666	cytochrome P450 2B
38	291.5	10.9	530	1	A27491	lanosterol 14alpha
39	290.5	10.8	502	2	T05246	cytochrome P450 mo
40	288.5	10.8	504	2	A29410	cytochrome P450, g
41	287.5	10.7	520	2	T24780	hypothetical prote
42	287	10.7	503	2	JC4702	cytochrome P450 3A
43	284.5	10.6	501	1	S38534	cytochrome P450 76
44	284	10.6	501	2	T04735	cytochrome P450 ho
45	282	10.5	491	2	T49625	testosterone 16a-h

## ALIGNMENTS

RESULT	1	
T46143	steroid 22-alpha-hydroxylase (DMF4) - Arabidopsis thaliana	
N:Alternate names: protein T3A5.40		
C:Species: Arabidopsis thaliana (mouse-ear cress)		
C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 16-Feb-2001		
C:Accession: T46143		
R:Bioecker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quettler, F.; S		
submitted to the Protein Sequence Database, December 1999		
A:Reference number: 223024		
A:Accession: T46143		
A>Status: preliminary		
A:Molecule type: DNA		
A:Residues: 1-513 <BL0>		
A:Cross-references: EMBL:AL132979		
A:Experimental source: cultivar Columbia; BAC clone T3A5		
C:Genetics:		
A:Map position: 3		
A:Introns: 74/2; 182/3; 233/3; 338/3; 369/3; 396/1; 432/3		
A>Note: T3A5.40		
C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology		
C:Keywords: heme; iron; metalloprotein		
F:308-484/domain: cytochrome P450 homology <P45>		
F:462/Binding site: heme iron (Cys) (axial ligand) #status predicted		
Query Match	99.9%; Score 2677; DB 2; Length 513;	
Best Local Similarity	99.8%; Pred. No. 9.2e-166;	
Matches	512; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MFETEHNTLLPLLPILLSLLLELLILKRRNRKTRFLPPGKSGWPLGTTGYLKYT 60	cytochrome P450 2B
DB	1 MFETEHNTLLPLLPILLSLLLELLILKRRNRKTRFLPPGKSGWPLGTTGYLKYT 60	testosterone 16alp
QY	61 ATTTGDEFMOQHVSKYKGRYSNLTGFEPTVSADGILNFILONGRLEPCYPSIGITL 120	cytochrome P450 2A
DB	61 ATTTGDEFMOQHVSKYKGRYSNLTGFEPTVSADGILNFILONGRLEPCYPSIGITL 120	cytochrome P450 2B
QY	121 GKMSLVLVGMHMDRMSISINFLSHARLRTLLKDYERHTLFVLDSDWQMSISADDEA 180	cytochrome P450 2B
DB	121 GKMSLVLVGMHMDRMSISINFLSHARLRTLLKDYERHTLFVLDSDWQMSISADDEA 180	cytochrome P450 2B
QY	181 KKFENFLMAKIMSDPGEEETDOLKKEYVTFPMKGVSAPIPLDGTAYHKAQSRATILK 240	cytochrome P450 2B
DB	181 KKFENFLMAKIMSDPGEEETDOLKKEYVTFPMKGVSAPIPLDGTAYHKAQSRATILK 240	cytochrome P450 2B
QY	241 FIERMEERKLDIKEDDEEVEVTEDEAEMSKSDHVKQRTDDDLGWLKHSNLSTEQ 300	cytochrome P450 2B
DB	241 FIERMEERKLDIKEDDEEVEVTEDEAEMSKSDHVKQRTDDDLGWLKHSNLSTEQ 300	cytochrome P450 2B
QY	301 ILDLISLLFGHETSVAIALAIFLQACRAVEELREERLEIARAKKEIGESELNMD 360	cytochrome P450 2B
DB	301 ILDLISLLFGHETSVAIALAIFLQACRAVEELREERLEIARAKKEIGESELNMD 360	testosterone 16a-h

Db 301 ILDLISLFLFAGHETSSVAIALAIFFLQACPAAVEELREELHETARAKKELGESELMWD 360  
QY 361 YKMDFTOCYINFTLRGNVRFLLHRRKALKDVRKYKGYDIPSGMKVLPVISAHLNDSRYD 420  
Db 361 YKMDFTOCYINFTLRGNVRFLLHRRKALKDVRKYKGYDIPSGMKVLPVISAHLNDSRID 420  
QY 421 QPMLFPMWMOQONNGASSSGSFSTWGNMYPFGGPRLCAGSELAKLEMAVFTIHLV 480  
Db 421 QPMLFPMWMOQONNGASSSGSFSTWGNMYPFGGPRLCAGSELAKLEMAVFTIHLV 480  
QY 481 LKFNWELAEODPFAFPVDFPGLPIRVSRIL 513  
Db 481 LKFNWELAEODPFAFPVDFPGLPIRVSRIL 513

## RESULT 2

S55379

cytochrome P450 CYP90 - Arabidopsis thaliana

N:Contains: oxidoreductase (EC 1.-.-.-)

C:Species: Arabidopsis thaliana (mouse-ear cress)

C.Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C.Accession: S55379

R:Sequences: M.; Nemeth, K.; Koncz, Z.; Nagy, F.; Koncz, C.

Submitted to the EMBL Data Library, May 1995

A:Reference number: S55379

A:Accession: S55379

A:Molecule type: mRNA

A:Residues: 1-472 &lt;S&gt;E&gt;

A:Cross-references: EMBL:X87367; NID:g853718; PIDN:CAA60793.1; PID:g853719

C:Genetics:

A:Gene: CYP90

C:Superfamily: Synechocystis cytochrome P450 slr0574; cytochrome P450 homology

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; monooxygenase;

F:775-440/Domain: cytochrome P450 homology &lt;P45&gt;

F:418/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 36.5%; Score 978.5; DB 1; Length 472;

Best Local Similarity 41.9%; Pred. No. 5.9e-56;

Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

QY 12 LLLPILSLFLILKRRNRKTRFNLPPGSGWPLGFTIGYLRKYATTLGDFMOOH 71  
Db 7 LLLLSLTAAGFL--LLRRTTRRRKMGSLPPSGSLPLIGTFOLIGYKTEPEPILDER 63  
QY 72 VSKYKRIYRNLGEPPIVSADAGLNKFLIQLNDGRLEFECSPRSIGILGKWSMLVVG 131  
Db 64 VARYGSVFMTLHGEPIFISADPETNRFLVQLNGKLFECSPASICNLGKHSLLMKGS 123  
QY 132 MHRDMSISLNFSLHARLITLKDVRHRLFLVDSMOONSIFSADDEAKKFFNLMAKH 191  
Db 124 LHRHMSLTMFSANSSITIKHMLDIDRLVFNLDSSSVLL--MEAKRITFEELTVKQ 181  
QY 192 IMNDGPEETDOLKREYVFMGVVASPLNPGTAVHKLQSRATILKIERKMEERKL 251  
Db 182 LMFDFG-EMSESLRKYELVIGFSLPLPLSTYTRKAIQAR-----KKVAELT 232  
QY 252 DIKEEDQEEBEVTEDEAEMSKSDHVRKQRTDGLGVLKHSNLSTEOILLILSLFA 311  
Db 233 VVVMKRRREEEGAE-----RKMDLAALLAADGGSDEEIVDFVALVA 278  
QY 312 GHETSSVAIALAIFFLQACPAAVEELREELHETARAKKELGESELMWDKMDFTQCVI 371  
Db 279 GYETSTIMTLAKFLTEFLAALQLEHEKIRAKSD--SYSLWSDYKSMPTQCV 336  
QY 372 NETLRLGNVRFLLHRRKALKDVRKYKGYDIPSGMKVLPVISAHLNDSRYDQPNLFPMRMQ 431  
Db 337 NETLRLGNVRFLLHRRKALKDVRKYKGYDIPSGMKVLPVISAHLNDSRYDQPNLFPMRMQ 431  
QY 432 QONNGASSSGSFSTWGNM--YMPFGGPRLCAGSELAKLEMAVFTIHLVLFKFNWELAE 490  
Db 397 -----SNSVTTGPSNVFTFPFGGPRLCAGSELAKLEMAVFTIHLVLTGFSWPAEQ 446

QY 491 DQFAFPVDFPGLPIRVSR 511  
Db 447 DKLVEFTTRTQKRYPLFVKR 467

## RESULT 3

H96759

probable steroid 22-alpha-hydroxylase T9L24.44 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C.Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C.Accession: H96759

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.M.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719

A:Accession: H96759

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-512 &lt;STO&gt;

A:Cross-references: GB:AE005173; NID:g11120803; PIDN:AG30983.1; GSPDB:GN00141

C:Genetics:

A:Gene: T9L24.44

A:Map position: 1

Query Match 32.3%; Score 866; DB 2; Length 512;

Best Local Similarity 33.8%; Pred. No. 1.2e-48;

Matches 181; Conservative 105; Mismatches 173; Indels 76; Gaps 7;

QY 12 LLLPILSLFLILKRRNRKTRFNLPPGSGWPLGFTIGYLRKYATTLGDFMOOH 66  
Db 11 LLSVSSSTFLFLITLLAGIARRRRRAHRLPPSSRGPLDGFVMAVAVAGSPSS 70  
QY 67 FMOOHVSK-----YGRYRNLGEPPIVSADAGLNKFLI 101  
Db 71 FVEKQIKKTFVSLCSVLLILKRPDNGFNEIRYKGLFSCSLGKAAVVSADDFRFRIM 130  
QY 102 QNDRLEFECSPRSIGILGKWSMLVVGDMHRDMSISLNFSLHARLITLKDVRHRT 161  
Db 131 QNDRLEFECSPRSIGILGKWSMLVVGDMHRDMSISLNFSLHARLITLKDVRHRT 161  
QY 162 LFLVDSMOONSIFSADDEAKKFFNLMAKHIMSDGEEETDOLKREYVFMGVVASPL 221  
Db 191 LQTLNFKDGEVVLADIDCKKVAIHLMVNOGLGS--SESEVDMQSOLFSDVDCLSVPI 249  
QY 222 NLPGTAVHKLQSRATILKIERKMEERKLIDKEEDQEEBEVTEDEAEMSKSDHVRKOR 281  
Db 250 DLPGFTYNNAMKARKKEIRINKTIERKLNKAAASD-----T 286  
QY 282 TDODLLGNVLKHSNLSTEOILLILSLRAGHETSSVAIALAIFFLQACPAAVEELREH 341  
Db 287 AGNGVGLRLEESLSPNESADPILNLFAGNETSTKTMFLFAVYFLHCPKATQOLLEH 346  
QY 342 LELARAKKELGESELMWDKMDFTQCVINFTLRGNVRFLLHRRKALKDVRKYKGYDIP 401  
Db 347 -----DRLAGMLTWQDYKTMFTQCVIYDELRLGGLIATLMLREKKEVSYQDIYIR 399  
QY 402 GMRVLPVISAHLNDSRYDQPNLFPMRMW-----QONNGASSSGSFSTWGNMYPFG 456  
Db 400 GCVVLPVISAHLNDSRYDQPNLFPMRMW-----QONNGASSSGSFSTWGNMYPFG 456  
QY 457 GCPRLCAGSELAKLEMAVFTIHLVLFKFNWELAEODPFAFPVDFPGLPIRVSR 511  
Db 450 GGTFRCPGALRIQIALFLHYFTYTKWTQLKEDRISFPSPARLVNGFKIQOLNR 504



RESULT 4  
D85429  
cytochrome P450 like protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 02-Apr-2001  
C:Accession: D85429  
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488  
A:Accession: D85429  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-457 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7270586; PIDN:CAB80304.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: At4g36380  
A:Map position: 4  
C:Superfamily: Synecocystis cytochrome P450 slr0574; cytochrome P450 homology  
C:Keywords: heme; iron; metalloprotein  
F:396/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 32.1%; Score 860; DB 2; Length 457;  
Best Local Similarity 36.8%; Pred. No. 2.6e-48;  
Matches 179; Conservative 93; Mismatches 160; Indels 54; Gaps 8;

QY 39 LPPKSGPFLGEGTIGY-----KPYATTLGDPMOOHVSKYKITYNSLFGPEITYSA 92  
D 2 IPNSLGLPVLGELTNFACGYSSRPVT-----FMCKRSLYGVFNFTNGIPITIST 55  
QY 93 DAGNRLFLONEGRLEECYSRISGILGKSMVLVGDMDRMSISLNFHARLRTI 152  
D 56 DAEVNVKVLQHGNTFYAYPKSITEELIGENSILSINGPHOKRLTLTGALRSRHLDR 115  
QY 153 LKQVRRHTFLVDSWQNSIFSAODEAKKFTFLMAKHISMDSPEEETOLKKEYVT 212  
D 116 ITRDIEASVNLTLASMAOLPLVHODEIKMTFELIVKVLSTSPG-EDMNLKLEFEF 174  
QY 213 MKGVASAPLNPGRAYHAKALOSRATILKFERKMEERLIDKEDEEVEEYKT-----E 266  
D 175 IKGICIPKIPKPRGLKYLKSLAKERLIKMKVKEERQVAMTTTSPADVVDVLRDGD 234  
QY 267 DEAEKMSDHYRKORTDDLLGWLKHNSLSTEQILLILSLFAGHETSSAIALATF 326  
D 235 SEKOSQSPDEYSGK-----ITVEMLPGEETMPTAMTLAYKF 270  
QY 327 LQACPRAVEELREHLEIARAKKEGESELMWDDYKKMDFTQCIVINETLRGNVRFPHR 386  
D 271 LSDNPVALAKLVEEMEMKRRKLEGE-EYKWTDYMSLFTQVNIETLRMANITINGYMR 329  
QY 387 KALDVRKGYDIPSGKVVLVISAVHLDNSRYDQNLFPNRMQOONNGASSSGSFS 446  
D 330 KALDVEKGLYIPKGMCVLASFSIVHDEDIYDNPYOFDMRMRIRINGSANSIC---- 385  
QY 447 TWGNVMPFGGPRCAGSELAKLEMAVEFHLVLFKNMELAEDDQPPAPFVDPNGLP 506  
D 386 -----FTFPGGGQRCLPGLIELSKLEISIFLHLVTRYSW-TAEDEIYSPFVKKRRLP 439  
QY 507 IRVSKI 512  
D 440 IRVATV 445

RESULT 5  
T48613  
hypothetical protein F18022.190 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T48613  
R:Beyan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24493

A:Accession: T48613  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-382 <BEV>  
A:Cross-references: EMBL:DB  
A:Experimental source: cultivar Columbia; BAC clone F18022  
C:Genetics:  
A:Map position: 5  
A:Insertions: 31/2; 139/3; 204/1; 224/3; 253/3; 280/1; 315/3; 349/2  
A>Note: F18022.190

Query Match 30.4%; Score 815; DB 2; Length 382;  
Best Local Similarity 34.2%; Pred. No. 1.7e-45;  
Matches 159; Conservative 87; Mismatches 127; Indels 92; Gaps 4;

QY 45 GMPFLGEGTIGYLYKPYATTLGDPMOOHVSKYKITYNSLFGPEITYSADAGINRFLONE 104  
D 2 GMPFLGEGTISFFKPHRSISGIFLQQRVSRYKVKRSNIGGKAVVSCDOELNMFILONE 61  
QY 105 GRLEPCSPRISIGILGKSMVLVGDMDRMSISLNFHARLRTILLKQVRRHTFLV 164  
D 62 GKLFSDYPRKAMHDLGYSLLATGELTRKLNKNTISFINLTSKPDPLCAENLSTSI 121  
QY 165 LDMQONSIFSAODEAKKFTFLMAKHISMDSPEEETOLKKEYVTPEMKGVASAPLNP 224  
D 122 LSKMKNCEVEFKEKVKLFTLSVAVNOLSTIKPEDRALVYLQDPLSTMKGFISLPLP 181  
QY 225 GTAYHAKALOSRATILKFERKMEERLIDKEDEEVEEYKTDEAEKMSDHYRKORTD 284  
D 182 GTGYNAIKVRSNRNIHONAIITEDMNNAIREDFDLSINSD----- 224  
QY 285 DLGVLNKHNSLSTEQILLILSLFAGHETSSAIALATFLOACPRAVEELREHLEI 344  
D 225 -----EEHNAI 230  
QY 345 ARAKKEGESELMWDDYKKMDFTQCIVINETLRGNVRFPHRKLKALDVRKGYDIPSGK 404  
D 231 -RAKKGDEL-LNMEDYKMEFTQCIVSEALRCGIVYTVHRAKTHDKFNEYVLPKGMK 288  
QY 405 VLVVISAVHLDNSRYDQNLFPNRMQOONNGASSSGSFSFTWGNVMPFGGPRCLAG 464  
D 289 VEPFTAVHLDPSLHENFEENPMPMKTKT-----AFGGGVRCVPG 329  
QY 465 SELAKLEMAVEFHLVLFKNMELAEDDQPPAPFVDPNGLPIRV 509  
D 330 GELGRLQIAFLHLHLVLSYRWKIKSDEMPVIAHPVEYFKRMGLEI 374

RESULT 6  
T07859  
cytochrome P450 homolog - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 16-Feb-2001  
C:Accession: T07859  
R:Bishop, G.J.; Harrison, K.; Jones, J.D.  
Plant Cell 8, 959-969, 1996  
A:Title: The tomato Dwarf gene isolated by heterologous transposon tagging encodes th  
A:Reference number: Z16181; MUID:96266705  
A:Accession: T07859  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-464 <BIS>  
A:Cross-references: EMBL:U54770; NID:g1421740; PIDN:AAI17070.1; PID:g1421741  
A:Experimental source: strain GCR/58  
C:Genetics:  
A:Gene: dwarf  
C:Superfamily: Synecocystis cytochrome P450 slr0574; cytochrome P450 homology  
F:273-436/Domain: cytochrome P450 homology <45>

Query Match 28.4%; Score 761.5; DB 2; Length 464;  
Best Local Similarity 33.4%; Pred. No. 6.2e-42;

Matches 170; Conservative 90; Mismatches 188; Indels 61; Gaps 10;

```

12 LLLPSLLSLFLILKLRN--RKTFRNLPKSGMPLGETIGLYKPTATTLG-DFM 68
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
5 LILSFFPGICITCALLRNNOVKYNOKNLPPTMGWPLFGETTEFLK-----LGP 58
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
69 QOHVSKYKTKYRNSLFGPEPTVSADAGLNRFILQNEGRLEFCSYPSISGIGLCKWMLVL 128
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
59 KNRARRGSEFKSHILGCPITVMSDELNRYILVNEAKGLVPGYPOSMIDILCKCINAAV 118
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
129 VGDHMDRMSISLNFSLHARLRTILKDVERRHLFVLDSQNSIFSADQAKFTFNLM 188
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
119 NSGAHYTMKALSLISPTWIRQDLPLKIDENRSHLTWN-DNKVIDIQETKMAFLSS 177
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
189 AKRIMSDPGEET---EOLKEVTFPMKGVSAPLNPGTAYHAKLQSRATILKEFERK 245
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
178 LKQI---AGIESTSLAQMSEFNLVLTSLPLNLPNTVNHRRGQAKIIVNLRL 233
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
246 MERKRLDIKEDEEVEVKTEDEAKMSKDVHVKORTDDLLGWLKHS---NLSTEQI 301
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
234 IEERR-----ASKETIOHMLGVLNNEATRKRLTDDDM 266
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
302 LDIILSLFAGHETSSVAIALAIFLQACPRAVEELREHLEIARAKKELGESELMWDY 361
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
267 IDILITILSGYTVSTTSMAYKYLHDPKYLEELRKEHMAIREKKP--EDPIDNDY 324
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
362 KKMDFQCVINETLRLGNVVRFLRRKALDVRKGYDIPSGMKVLPVISAHLNDSRYDQ 421
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
325 RSMFRVAVILETSLATVINGVLRKTQDMETNGYIIPGMRIVYVTRLENDPRLYPD 384
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
422 PNFNPMRMOQONNGASSSSGFSFTWGNVMPFGGPRICASELAKLEMAVFIHHLVL 481
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
385 PYSFNPMRWMDKS-----LEHNSFLVFGGTRQCPKEKELGVAISFLHYEVT 433
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
482 KFNWELAEDDQPPAPFVDFPNGLPIRVS 510
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
434 KYRWEELGGDKLMKPFVPEARNGLRIRVS 462
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

# RESULT 7

T04444  
 cytochrome P450 - Arabidopsis thaliana  
 N:Alternate names: protein T18B16.200; protein T5K18.10  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 28-Jul-2000  
 C:Accession: T04444; T05806  
 R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Bancroft, I.; Mewes, H.W.  
 submitted to the Protein Sequence Database, April 1998  
 A:Reference number: Z15359  
 A:Molecule type: DNA  
 A:Accession: T04444  
 A:Residues: 1-457 <BEV>  
 A:Experimental source: EMBL:AL021687  
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voec, M.; Robben, J.; Volckaert, G.; Be  
 submitted to the Protein Sequence Database, April 1998  
 A:Reference number: Z15453  
 A:Molecule type: DNA  
 A:Residues: 131-457 <BEW>  
 A:Cross-references: EMBL:AL022580  
 A:Experimental source: cultivar Columbia; BAC clone T5K18  
 C:Genetics:  
 A:Map position: 4  
 A:Introns: 67/2; 173/3; 302/3; 358/1; 393/3  
 A:Note: T18B16.200; T5K18.10  
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology  
 C:Keywords: chromoprotein; electron transfer; heme; iron; monooxygenase; oxidoreductase  
 F:272-433/Domain: cytochrome P450 homology <P45>

Query Match 23.6%; Score 632; DB 2; Length 457;

Best Local Similarity 31.8%; Pred. No. 1,5e-33;  
 Matches 151; Conservative 92; Mismatches 176; Indels 56; Gaps 12;

```

13 LLLPSLLSLFLILKLRNRRKTRFNLPPKSGMPLGETIGLYKPTATTLGDFM 72
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
10 LFAQSLFLFLRCLISQRRGSKLPLPPTMGWPLFGET---FQLSQDP-NVFFQSKO 65
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
73 SKYKTYRNSLFGPEPTVSADAGLNRFILQNEGRLEFCSYPSISGIGLCKWMLVLGDM 132
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
66 KRYGVFKTHVLGCPVMISSPEAKFVLVTKSHLRFKPFPAKREMLGKQALFFHQGDY 125
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
133 HRDMSISLNFSLHARLRTILKDVERRHLFVLDSQNSIFSADQAKFTFNLAHNI 192
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
126 HAKRLVLRAPMESIRN-MVPIESIAODSLRSG-EGMINTYQEMTYFFNVALIST 183
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
193 MSMDPGEET---EOLKEVTFPMKGVSAPLNPGTAYHAKLQSRATILKEFERMEER 249
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
184 F---GKDEVLVEDLKRCLYILEKGYNSMPVLPBTLFRKSKAKAKELSQLARLSER 239
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
250 KLDIKEDEEVEVKTEDEAKMSKDVHVKORTDDLLG-WVLKHSNLSTEQILDLISL 308
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
240 R-----QNGSSH-----NDLGSFMGDKBELDQIADNIGV 272
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
309 LFAGHETSSVAIALAIFLQACPRAVEELREHLEIARAKKELGESELMWDYKKMDFQ 368
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
273 IFARDPTASVMSWILKYLAENPNVLEAVEEQMAI-RDKEGES-LTWGDTKKMPLTS 330
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
369 CVINETLRLGNVVRFLRRKALDVRKGYDIPSGMKVLPVISAHLNDSRYDQPNLFNPY 428
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
331 RVIGETLRVASILSFTFREAVEDEVEGYILIPKGMVLPFLFRHHSADIFSHPGDFPS 390
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
429 RMOQONNGASSSSGFSFTWGNVMPFGGPRICASELAKLEMAVFIHHLVKEF 483
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
391 RFE-----VAKPNTFMFPGNGTHSCPNELAKLEMSIMIHLLTKY 432
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

# RESULT 8

T02739  
 probable cytochrome P450 At2g29090 [imported] - Arabidopsis thaliana  
 N:Alternate names: cytochrome P450 homolog T914.17  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Feb-2001  
 C:Accession: T02739; D84692  
 R:Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.;  
 submitted to the EMBL Data Library, August 1998  
 A:Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.  
 A:Reference number: Z14710  
 A:Accession: T02739  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-482 <ROU>  
 A:Cross-references: EMBL:AC005315; NID:g3461834; PID:g3461849  
 A:Experimental source: cultivar Columbia  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,  
 eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: D84692  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-482 <STO>  
 A:Cross-references: GB:AE002093; NID:g3461849; PIDN:AAC33235.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g29090; T914.17  
 A:Map position: 2  
 A:Introns: 80/2; 187/3; 321/3; 351/3; 413/3; 453/2  
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology  
 F:291-453/Domain: cytochrome P450 homology <P45>

Query Match 22.8%; Score 611; DB 2; Length 482;

Best Local Similarity 29.88; Pred. No. 3.5e-32;  
Matches 153; Conservative 101; Mismatches 192; Indels 68; Gaps 13

Oy 16 PSLISLILFLILK-----RRNRKRENLPGSGMFEELGTYLCKPYATLILGDMQ 69  
 Db 20 PALITLTVVVVVVLLFEKWLTHWKEQRLRLPPSGKGLPIGET--LRLYENP-NSFEA 75  
 Oy 70 QHVSRYGKIVYSNLFGEPTIVSADAGLNRFTLIONEGRLEFECSPYRSIGLIGKSMVLV 129  
 Db 76 TRONKYGDIIFPHILIGCCVWIMSSPEARWLVLSAHLKFTYPPSKRMIGPEALPHQ 135  
 Oy 130 GDMHDM-RSISLNFLSHARLRTILKDKVERHTLVFVLDWOONSIFSADQAKFTENLM 188  
 Db 136 GPYHSTLRLVSSFPSPALRPTV--SHIELLVLTQTLSSMTSOKSINTLEMKRYAEFVA 193  
 Oy 189 AKHIMSMDPGE---TBOLKKEYTFPMKGVVSAPLNLPGAYKALQSRATILKFIERK 245  
 Db 194 ---LMSAFDGEELPTTIDVILKTLVRLERGSYMSPLDGLFLPHSKMARTELESEELKV 250  
 Oy 246 MEERLIDKEEDOEVEEVEKTEDEAEMSKSDVHRKORTDDLLGWLV----KSNLSTEQ 300  
 Db 251 IEKRENRRE-----GGLLVGLGAKDKRNLDSQ 283  
 Oy 301 ILDLISLLEPGHETSSVAIALAIFLOACPRAVEELREHLEI-ARAKKELGSELNWD 359  
 Db 284 IADNIGIVFATPDTTASVLTLWLKLYLHDHNLLOEVSREOFSTOKIKKE--NRRISWE 341  
 Oy 360 DYKKMDPQCVINETLRGNVVRFLRKALDVRKGVDPISGMVPLVISAHLNDNRY 419  
 Db 342 DTRKMLTTRVIOELTRASVLSFTFRVADVDYEYDGLILIRKGMVLPFLRRIHSSSEF 401  
 Oy 420 DQPLNLFNWRMOOONNGASSSGSFSTWGNVMPFGGPRLCAGSELAKTEMAVFIHL 479  
 Db 402 PDPEKFDSPRFE-----VAPKPYTYMGFGVNSCPSELAKIEMILLHL 448  
 Oy 480 VLKFWELAEDDQPPAF-PFVDFPNGLPPIRYSRI 512  
 Db 449 TTSFRWEVIGDEEGIQYGPFPVKRGLPIRVTP 482  
  
 RESULT 9  
 C71417  
 cytochrome P450 d13695c - Arabidopsis thaliana  
 N:Contains: oxidoreductase (EC 1.-.-.)  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 A:Variety: columbia  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
 C:Accession: C71417  
 R:Byman, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk  
 R:Wedler, H.; Wedler, E.; Wandbut, R.; Weltzienegger, T.; Pohl, T.M.; Terry, N.; Giel  
 R:vanagah, T.; Hempel, S.; Kotter, P.; Entlian, K.O.; Rieger, R.; Schaeffer, M.; Funk, B.  
 R:Nature 391, 485-488, 1998  
 A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Mortfort, A.; Pons, A.; Puigdomenech  
 R:erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reehman, S.; Anse  
 R:C.; Chalwatzis, N.  
 A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha  
 A:Reference number: A71400; MUID:98121113  
 A:Accession: C71417  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1487 <BEV>  
 A:Cross-references: GB:297338; NID:g22444870; PIDN:CAB10309.1; PID:g22444888  
 C:Genetics:  
 A:Gene: d13695c  
 A:Map position: 4COP9-4G3845  
 C:Superfamily: Synchocystis cytochrome P450 slt0574; cytochrome P450 homology  
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
 E7433/Binding site: heme iron (Cys) (axial ligand) #status predicted

[illegible]

Query Match	21.7%	Score 581;	DB 1:	Length 487;
Best local Similarity	27.3%	Pred. No.	3.1e-30;	
Matches 143;	Conservative 100;	Mismatches 205;	Indels 76;	Gaps 11;

  

OY	65	GDPMQAHVSKGKIRSNLFEPTIVSDAGLNRFILIONERLEFECSPRISGLIKWS	12
:     :   :	:	:	:  :

Db 72 EDVNPRIKIGNTFKRIRNGSPITVNGAEANRLIISNEFSLVSSWSPSSVOLMGNC 131

Qy 125 MLVLGDMHDMRSISLNFSLHARLTLLKDVHHTLFLVDSMOQNSIFSADAEAKFT 184

Db 132 IMAGQEKHRLRGIVANSLSYIGLESILPKLCPTVKPHHTEMRGKREISLYSAKYL 191

Qy 185 FNLMAKHL--MSMDGEETBOLKKEYTEPMKGVASAPLNPATYAKALOSRATILKFI 242

Db 192 FTVVEECYLGKIVETGMLEV-----FERYLEGVAFALVEPEPCSKFARAKRAREIEFTL 245

Qy 243 ERKMEERKLDKEEDOEDEEVEKTEDEAEMSKDVHVKORTDOLLGWLKHSNSTEOL 302

Db 246 VGKREKREKREKREKAE-----KPNLTLSRLVEELIKGV-----ITEEV 287

Qy 303 DLILSLFAGHETSSVALAIAIFLQACPKAVEELREEHLEIARAKKEGESE-LMMDY 361

Db 288 DNMVLVFAADHTSYAASMTFMKLAQHPRCDRLDLEHAQI---KANKGEXELTVEDV 344

Qy 362 KKMFTQCVINETRLGNVAVFLHRKALKDVRKYGYDIPSGWKVLPVISAHLNDSRYDQ 421

Db 345 KKMYSQVQVETRWKLSPPRIGSFRKAVADIDGGYTIIPKGMKILMTYTGTHYPELFQD 404

Qy 422 PNLFPNPMQOONNGASSSGSFSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHLVL 481

Db 405 PMSDFPRFDFKPIQAYT-----YLPFGGGRCLCAGHQLAKISLIVFMHEVYT 451

Qy 482 KFNWELAEDDQPFAPFVDFPN-GLPIRVS 510

Db 452 GFDMSLVYPRDETISMDFPPLGMPKIRIS 481

## RESULT 11

H86185  
hypothetical protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: H86185  
R:Phenologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Hultzer, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: H86185  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-490 <STO>  
A:Cross-references: GB:AE005172; MID:g2388581; PIDN:AAB71462.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1  
C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology

Query Match 21.6%; Score 579.5; DB 2; Length 490;  
Best Local Similarity 28.6%; Pred. No. 4e-30;  
Matches 133; Conservative 107; Mismatches 192; Indels 33; Gaps 9;

Qy 34 KTRNLPKSGMGPFLGFTIGLKPYATTLGDFMOQHVSKYK--IYRSNLFSEPTIVS 91

Db 40 ENRYHLPKPGDGLPFGPILGMSLFLRAFKTSDPSFTRLIKRYGPKGIYKAMFENPSTIV 99

Qy 92 ADAGLNRFILONEGRLEPCSPRSIGILGKWSMLVVGDMHDMRSISLNFSLHARLT 151

Db 100 TTSOTCRVNLVDD--AFKPGWPTSMELIGKSKSVGVSFEHKKRLRLTAAPVNGHEALS 158

Qy 152 ILKDVHHTLFLVDSMOQNSIFSADAEAKFTFNLMAKHIMSDPGEETEOLKKEYVT 211

Db 159 TYIYIEENVTIVLDKMTKMGEEFELHRLKRLTRIT-MYILSSESSENVDALEREYTA 217

Qy 212 FMKGVASAPLNPSTAYAKALOSRATILKFIERKMEERKLDKEEDOEDEEVEKTEDEAEM 271

Db 218 LNVGVAAMAVNIPGAVYHARLAKARTLVAAQSLYTER-----NQRKQIILSKKQML 271

Qy 272 SKSDHYKQKRTDOLLGWLKHSNSTEOLDLILSLFAGHETSSVALAIAIFLQACP 331

Db 272 DNLNIVK-----DED-----GKTLDEEITIDVILMYLMAGHSSGHTIMATVFLQENP 320

Qy 332 KAVEELREEHLEIARAKKEGESELMMDYKKMFTQCVINETRLGNVAVFLHRKALKD 391

Db 321 EYLQAKAEQMLIKSRE--GQKLSLKETKMEFLSVYDETRIVITFSLTAREAKTD 379

Qy 392 VRYKYDIPSGWKVLPVISAHLNDSRYDQNLFPNPMQOONNGASSSGSFSFTWGNM 451

Db 380 VEMNGYLLPKGMKVLGFERDYNHPEVPRDKKDPARW---DNG-----FVPRAGA 428

Qy 452 YMPFGGPRCLCAGSELAKLEMAVFIHLVLKFNWELAEDDQPFAP 496

Db 429 FLPRGAGSHLCPGNDLAKLETISFLHHLKTYQVKSNSPECPVM 473

## RESULT 12

cytochrome P450-like protein - Arabidopsis thaliana  
N:Alternate names: protein F14D17.40  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
C:Accession: T48973  
R:Jordan, N.; Bangert, S.; Widelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: 225008  
A:Accession: T48973  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-455 <JOR>  
A:Cross-references: EMBL:AL353992; GSPDB:GN00061; ATSP:F14D17.40  
A:Experimental source: cultivar Columbia; Bac clone F14D17  
C:Genetics:  
A:Gene: ATSP:F14D17.40  
A:Map position: 3  
A:Insertions: 70/2; 178/3; 229/3; 312/3; 343/3; 346/1; 381/3  
C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology

Query Match 21.5%; Score 577; DB 2; Length 455;  
Best Local Similarity 28.9%; Pred. No. 5.2e-30;  
Matches 145; Conservative 84; Mismatches 173; Indels 100; Gaps 10;

Qy 32 NKRTRNLPKSGMGPFLGFTIGLKPYATTLGDFMOQHVSKYKTYRSNLFSEPTIVS 91

Db 28 NPKSNGKLPKSGMGPFLIGETLDFKPYGEYEISPYLKKMLKRLPFRITNLGVKTVVS 87

Qy 92 ADAGLNRFILONEGRLEPCSPRSIGILGKWSMLVVGDMHDMRSISLNFSLHARLT 151

Db 88 TDKQVNMELIOENKSFILSTPDGIMKFLKDSLFKIGNHKKIKQITLLLSSEGILK 147

Qy 152 ILKDVHHTLFLVDSMOQNSIFSADAEAKFTFNLMAKHIMSDPGEETEOLKKEYVT 211

Db 148 KILKDMQVREHLSNAKTRGLDVKAIVSKLIAHLTPKMS-----NLKPTQA 198

Qy 212 FMKV-----VASPLNPSTAYAKALOSRATILK--FIERKMEERKLDKE 255

Db 199 KLMGIFRAFTFDMERTSYLSAGGLVNTLM--ACRGMREIKIDYIMRKSEKY---- 252

Qy 256 EDGEEVEKTEDEAEMSKDVHVKQRTDOLLGWLKHSN-----LSTEOLDLILSLF 310

Db 253 -----DPLNLAIESEKAGELNDAITLITFLSC 284

Qy 311 AGHETSSVALAIAIFLQACPKAVEELREEHLEIARAKKEGESELMMDYK-KMFTQ 369

Db 285 VTQDTTKALICLAVKFLLENKVLVLAELKKEH-EVILSRDEKBCGVYMEVRYHNMFTFN- 342



T02263  
 cytochrome P450 DWARF3 - maize  
 N:Contains: oxidoreductase (EC 1.-.-.-)  
 C:Species: Zea mays (maize)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: T02263  
 R:Winkler, R.G.; Helentjaris, T.  
 Plant Cell 7, 1307-1317, 1995  
 A:Title: The maize dwarf3 gene encodes a cytochrome P450-mediated early step in gibberellin biosynthesis  
 A:Reference number: 214648; MUID:96004534  
 A:Accession: T02263  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-519 <WIN>  
 A:Cross-references: EMBL:U32579; NID:9987266; PIDN:AAC49067.1; PID:9987267  
 A:Experimental source: strain B73  
 C:Genetics:  
 A:Gene: dwarf3  
 C:Function:  
 A:Description: Involved in an early step in gibberellin biosynthesis  
 A:Pathway: gibberellin biosynthesis  
 C:Superfamily: Synchocystis cytochrome P450 slr0574; cytochrome P450 homology  
 C:Keywords: oxidoreductase  
 F:325-488/Domain: cytochrome P450 homology <P45>

Query Match 18.3%; Score 491.5; DB 1; Length 519;  
 Best Local Similarity 28.2%; Pred. No. 2,1e-24;  
 Matches 129; Conservative 91; Mismatches 189; Indels 49; Gaps 11;

QY 36 RFNLPPGKSGMPLGEGTGYKPYTATLGDPMOHWKSKYK--TYSNLFGEPTI-VSA 92  
 Db 69 RARLPPEMGWPLVGGWMAFRAKSGKPDFAFVFRFRGTGVYRSFMSSPTLVTT 128  
 QY 93 DAGNRFILONEGRLEPCSYRSTIGILGKWSMLVLVGDMDRMSISLNFSLHRLRTI 152  
 Db 129 AEGCKQVLMDDA--FVTGMPKATVALGPRSFYAMPYDEHRLRLKLTAPDINGFDALTG 186  
 QY 153 LKLDVERHTLEFLDSW-QQNSIFSADDEAKKFTENLMAKHIMSMDPGEETEQLKKEYVT 211  
 Db 187 YLPFIDRTVTSLSRAMADHGGSVEFTELRRMTEKTIYQ-IFLGADQATRALERSYTE 245  
 QY 212 FMKGVVASPLNLPGTAYHKKALQSRAITLTKTERKMEERKLDIKE-----EDQEEVEYKT 265  
 Db 246 LNYGMRAMAINLPGFAYRGALRRARRLVAVLQGLDERRAARAKVSGGVDMDRLIEA 305  
 QY 266 EDEAEMSKSDHVRKQRTDDDLGWLKHSNLSSTQILDILISLFAGHETSSVALALAI 325  
 Db 306 QDE-----RGRHDDDD-----EITDVLVMTLNAGHSSGHTMTMATV 342  
 QY 326 FLQACPKAVEELREHELEIARAKKELGESELNMDYRKMDFTQCVINETLRLGNVRFLLH 385  
 Db 343 FLQENPDMFAFAKAEQALMRSIPS-SORGLTLNDFRMEYLSQVIDETLRLVNISSVSF 401  
 QY 386 RKALKDVRKYCYDLPISGKYLPIVISAHLNLSRYDQNLNFMWQOONNGASSSGSGSF 445  
 Db 402 RQATRDVFNQGYLLPKGMKQVLMYRSVAMDPQVYDPDKFDPSSWE----GHSPPAGTF 456  
 QY 446 STWGNMYPEGGGRLCAGSELAKLEMAVFIHLVLKF 483  
 Db 457 -----LARGLGARLCPGNDLAKLEISVFLHFLG 487

Search completed: September 28, 2001, 18:45:05  
 Job time: 95 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 : Search time 14.33 Seconds  
(without alignments)  
737.114 Million cell updates/sec

Title: US-09-502-426-2

Perfect score: 2681  
Sequence: 1 MFEHTHTLLPLLLPLSLLS.....FAFPVDFPGLPIRYSRL 513

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	978.5	36.5	472	2	US-08-622-166A-2
2	978.5	36.5	472	2	US-08-622-166A-4
3	503	18.8	492	3	US-08-724-466B-2
4	439	16.4	497	3	US-08-724-466B-4
5	289	10.8	511	4	US-08-991-677-4
6	279.5	10.4	513	3	US-08-948-564-6
7	264	9.8	520	2	US-09-091-432-2
8	263.5	9.8	504	1	US-08-457-274A-25
9	263.5	9.8	504	5	PCT-US95-05758-25
10	257	9.6	476	1	US-08-313-075A-30
11	256.5	9.6	490	1	US-08-201-118-7
12	256.5	9.6	490	2	US-08-238-821B-7
13	256.5	9.6	490	5	PCT-US95-05744-7
14	245	9.1	495	1	US-08-532-065B-2
15	244.5	9.1	576	3	US-08-948-564-16
16	241	9.0	510	3	US-08-606-505B-66
17	241	9.0	510	4	US-09-616-990-66
18	240.5	9.0	509	3	US-08-948-564-18
19	239	8.9	496	1	US-08-313-075A-50
20	236	8.8	521	1	US-08-948-564-14
21	234.5	8.7	490	1	US-08-201-118-1
22	234.5	8.7	490	2	US-08-238-821B-1
23	234.5	8.7	490	4	PCT-US95-05744-1
24	233.5	8.7	508	4	US-08-991-677-2
25	231	8.6	506	1	US-08-313-075A-38
26	231	8.6	506	3	US-08-606-505B-65
27	231	8.6	506	4	US-09-616-990-65

28	223.5	8.3	496	4	US-09-292-768-64	Sequence 64, Appl
29	223	8.3	490	1	US-08-201-118-3	Sequence 3, Appl
30	223	8.3	490	2	US-08-238-821B-3	Sequence 3, Appl
31	223	8.3	490	5	PCT-US95-05744-3	Sequence 9, Appl
32	222	8.3	490	1	US-08-201-118-9	Sequence 9, Appl
33	222	8.3	490	2	US-08-238-821B-9	Sequence 9, Appl
34	222	8.3	490	5	PCT-US95-05744-9	Sequence 9, Appl
35	222	8.3	516	3	US-08-948-564-12	Sequence 12, Appl
36	221.5	8.3	500	2	US-08-314-601-2	Sequence 2, Appl
37	221.5	8.3	500	5	PCT-US95-13051-2	Sequence 2, Appl
38	218	8.1	510	3	US-08-948-564-4	Sequence 4, Appl
39	215	8.0	500	3	US-08-881-784-9	Sequence 9, Appl
40	214.5	8.0	496	3	US-08-881-784-1	Sequence 1, Appl
41	214.5	8.0	496	4	US-09-292-768-2	Sequence 2, Appl
42	214.5	8.0	496	4	US-09-292-768-66	Sequence 66, Appl
43	214	8.0	490	1	US-08-201-118-13	Sequence 13, Appl
44	214	8.0	490	2	US-08-238-821B-13	Sequence 13, Appl
45	214	8.0	490	5	PCT-US95-05744-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-622-166A-2  
Sequence 2, Application US/08622166A  
Patent No. 5952545  
GENERAL INFORMATION:  
APPLICANT: KONZ, CSABA  
APPLICANT: MATUR, JAIDEEP  
APPLICANT: SZEKERES, MIKLOS  
APPLICANT: ALTMANN, THOMAS  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING CYTOCHROME  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Birch, Stewart, Kolasch & Birch, LLP  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/622,166A  
FILING DATE: 27-MAR-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: SVENSSON, LEONARD R.  
REGISTRATION NUMBER: 30,330  
REFERENCE/DOCKET NUMBER: 0147-0153P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 205-8000  
TELEFAX: (703) 205-8050  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 472 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-622-166A-2

Query Match 36.5%, Score 978.5, DB 2, Length 472:  
Best Local Similarity 41.9%, Pred. No. 21e-89,  
Matches 210, Conservative 171, Indels 41, Gaps 8,  
QY 12 LLLPLSLILFLILKKRRNRKTRFNLPPGKSGPFLGFTIGYILKPYATVATLGDPMOQH 71





APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 492 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-724-466B-2

Query Match 18.8%; Score 503; DB 3; Length 492;  
Best Local Similarity 27.6%; Pred. No. 8.4e-42;  
Matches 144; Conservative 95; Mismatches 216; Indels 66; Gaps 14;

QY 7 HTLPRLPLSLSLLEL-----ILKRRNRKTRFNPFGSGWPFGLGTGYLKP 58  
DB 4 YLMTFTICTIVLPVLLFAAVKLEMEMIMIRRVDPNCRSLPGLPTMGFLPFIETQLI-- 61  
QY 59 YVATLPGPMOQHVSKYGIYRSNLPGEPTIVSADAGLNRFILONEGRIFECYSRISIG 118  
DB 62 ---LORRFLNRKRKQYCGITVTHLFGNPTVYMGADNVKQLLGEHKIVSYQWPAVST 118  
QY 119 ILGKSMVLVVGDMHRD-----MRSISLNFSLSHARLRTLLKDVERTHFLVDSW-QONS 172  
DB 119 ILGSLTSLNVHGVQKNNKKAIMRAFNRDLRH-----YIVIQEVKSAIOENLQKNS 172  
QY 173 IFSADDEAKTFPNNMAKIHMSDPEEET--EQLKKEYVTMKGVASAPLPGTAHYK 230  
DB 173 CVLVYPEMKKIFRIAMRLLLGEPQLITDEQLVEAEEMIKNLSLPIDVPSGLYR 232  
QY 231 ALQSRATILKFERMEERKLDIKEDEEVEKTEDEAEKMSDHYVRQRTDDDLGV 290  
DB 233 GLRAR---NFIHSTIEIRKKIODDNEOKYKDALQLL-----TENSRSDE----- 279  
QY 291 LKHSNLSTEQILDLILSLFAGHETSSVAIALAIFLQACPAVEELREHELEIRAKKE 350  
DB 280 ---PFSIQAMKEATFELFGHETTASTATSLVWELGINTVEYQVRE-----VQEKYE 331  
QY 351 LG---ESELNDDYKKMDFTOCVINEIRLGNVYFELRKALKDVRKYGYDIPSGKYL 406  
DB 332 MGMTYPGKLSMELLDOLKYTCVIKETLRINPVPYCGFRVALKTFELNGYOIPKGMVY 391  
QY 407 PIVSAVHLDSRYDQPNLFNPRMOQONNGASSSGSSTGMNNYMPGGGPRCLAGSE 466  
DB 392 YSICDTHDAVDFPNKEEFQPERFM-----SKGLEDSRP--NZIPGGGSRMCVKE 442  
QY 467 LAKLEMAVEFIHLVLKFNWELAEDDQ-----PFAFPEYDFP 502  
DB 443 FAKVLKIFLVELLQHCNMIISNGPPTMTKGTPIYVVDLP 483

RESULT 4  
US-08-724-466B-4  
Sequence 4, Application US/08724466B  
Patent No. 6063606  
GENERAL INFORMATION:  
APPLICANT: Belkovich, P. Martin, White, Jay A.,  
APPLICANT: Beckett, Barbara R., Jones, Glenville  
TITLE OF INVENTION: Retinoid Metabolizing Protein  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
ZIP: M5L 1A9

COUNTRY: Canada  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 MB storage  
COMPUTER: COMPAQ, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,466B  
FILING DATE: October 1, 1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/667,546  
FILING DATE: June 21, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 50767/00004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 497 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-724-466B-4

Query Match 16.4%; Score 439; DB 3; Length 497;  
Best Local Similarity 27.0%; Pred. No. 2.2e-35;  
Matches 142; Conservative 85; Mismatches 226; Indels 72; Gaps 15;

QY 10 LPLLLPLSL-----LSLLELILLL-----RRNRKTRFNPFGSGWPFGLGTGYL 57  
DB 3 LPLALSALCTVLPVLLFLAIVKIMDLVYCSGRDSCALPLPGTGFPGFGETL----- 58  
QY 58 PYATLPGPMOQHVSKYGIYRSNLPGEPTIVSADAGLNRFILONEGRIFECYSRISIG 117  
DB 59 -QWVLRKRFLOMKRRKQYFIYKTHLFGNPTVYMGADNVKQLLGEHKKIVSYQWPAVST 117  
QY 118 GLIGKSMVLVVGDMHRDMSISLNFSLSHARLRTLLKDVERTHFLVDSWQO--NS 172  
DB 118 TLGSGCLSNLHDSHQRKQYIMRAFNRDLRH-----LEQWLSGGER 173  
QY 173 IFSADDEAKTFPNNMAKIHMSDPEEET--EQLKKEYVTMKGVASAPLPGTAHYK 227  
DB 174 GLLVYPEVNRKIFRIAMRLLLGEPQLAGDQSE-QQLVEAEEMTRNLSLPIDVPSG 232  
QY 228 YHKAQSRATILKFERMEERKLDIKEED-----OEEVEKTEDEAEKMSDHYVRQRT 282  
DB 233 LYRGKARLNLIRIQRNIRAKICGRASEAGCCDALQLLIEHSE-----RGERL 285  
QY 283 DDQLGLVNLKHSNLSTEQILDLILSLFAGHETSSVAIALAIFLQACPAVEELREHL 342  
DB 286 DMQ---ALKQ---STE-----LLFGHETTASTATSLVWELGINTVEYQVREELK 331  
QY 343 ELARAKKEGESELMNDYKKMDFTOCVINEIRLGNVYFELRKALKDVRKYGYDIPSG 402  
DB 332 SKGLCKSNQDKLDMELIOLKYTCVIKETLRINPVPYCGFRVALKTFELNGYOIPK 391  
QY 403 WKVLPVSAVHLDSRYDQPNLFNPRMOQONNGASSSGSSTGMNNYMPGGGPRCLAG 462  
DB 392 WNVYISICDTHDAVDFPNKEEFNPRESAPHPEDASRFS-----FIPFGGGLRSC 442  
QY 463 AGSELAKEMAVEFIHLVLKFNWELAE-----DDQFAPPEYDFP 502  
DB 443 VGEKEFAKILKIFVELARHCWQMLNGPPTMTKGTSTYVYVVDLP 487

RESULT 5  
US-08-991-677-4  
Sequence 4, Application US/08991677A  
Patent No. 6252135

```

: GENERAL INFORMATION:
: APPLICANT: Chiang, Vincent L
: APPLICANT: Carraway, Daniel T
: APPLICANT: Smeltzer, Richard H
: TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
: FILE REFERENCE: 50617
: CURRENT APPLICATION NUMBER: US/08/991.677A
: CURRENT FILING DATE: 1997-12-16
: EARLIER APPLICATION NUMBER: US 60/033,381
: EARLIER FILING DATE: 1996-12-16
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 511
: TYPE: PRT
: ORGANISM: Liquidambar styraciflua
US-08-991-677-4

Query Match          10.8%; Score 289; DB 4; Length 511;
Best Local Similarity 22.0%; Pred. No. 2.3e-20;
Matches 121; Conservative 95; Mismatches 196; Indels 138; Gaps 21;

QY 6 HTHTLPL-LILPSLILFLILKRRNRKTRFNLPPKSGMPFLGETIGYLTATTL 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 6 HEALQPLMTLFTIPLILLLGLVSR--LRRLPYPEPKGLPIVGNML--MDQLTHRGL 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 65 GDFMOHVSXKYKIRSNLFEPTIVSADAGLNRFILNLSGRLE-----CSYPR 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 AKLAKQ-----YGLFLHLMGFLHVVAVSTPRMAROVLOVQONIFSNRATIAISLTYYDR 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 SIGGILGKWSMLVVGDMHNRMSIS--LNFSLHARLRTILKDYERHTLFLVDSMOONSI 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 119 AD-----MAFAHYGPEFRQMRKLCVAKLFSSRR--AESME----- 151
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 174 FSAODE-----AKKFTFNLMAKHIMSMDPGE-----EPTTE 203
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 152 -SVNDEVDASAVRVASNGSTVINGELVFALTKNITTRAAGTISHEDQEFVAILOEFS 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 204 QL-----KREYVTEMKGV--VSAPLNLPGTAHKALOSRAATILKFTERKMEERKLDIKE 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 211 QLFQAFNLADRIPLMLKNVPOGINVRLN-----KARGALDGFIDKIIDH---IQK 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 EDQEEERYKTE-----DEAEMSKSDHVRKQRTDDLLGLVNLKHSNSTQOIIDLT 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 258 GSKSSEEDVDMDLLAFYGESEAKVSESD-----LONSILKTDNIDKAI 303
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 306 LSLFAGHETSSVALAIFLQACPAVEELREEHLEIARAKKEGSELNMDYKMD 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 304 MDVAFGGETVASAIEAMTEFLKSPEDLKVKQGLAVVVGDLDRVEEK-----DFEKLIT 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 366 FTQCVINETLRAGNVVREFLHRAKALKDVRKYKGYDIPSGMKVLPVTSAVHLDSRYDQPNLF 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 359 YLKVCLKEVLRHPRIPILLHETAEDEVGYYIIPAKSRVINACAIGRDKNSSADPDTE 418
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 426 NPMWMOONNGASSSGSFSFTWGN--YMPFGGPRLCAGSELAKLEMAFIHHLVLEKF 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 419 RPSFELK-----DGYPDRK--GNNEEFIFGSGRSCPEMOLGYALLETVAHLHLCF 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 484 NMELAEEDQD 493
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 470 TWELPDGMR 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-08-948-564-6
: Sequence 6, Application US/08948564
: Patent No. 6121512
: GENERAL INFORMATION:
: APPLICANT: Siminszky, Balazs
: APPLICANT: Dewey, Ralph E.
: APPLICANT: Cordin, Frederick T.
: TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
```

```

: TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Virginia C. Bennett
: STREET: PO Box 37428
: CITY: Raleigh
: STATE: No. 6121512th Carolina
: COUNTRY: USA
: ZIP: 27627
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/948,564
: FILING DATE:
: CLASSIFICATION: 800
: ATTORNEY/AGENT INFORMATION:
: NAME: Bennett, Virginia C.
: REGISTRATION NUMBER: 37,092
: REFERENCE/DOCKET NUMBER: 5051-409
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 919-854-1400
: TELEFAX: 919-854-1401
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 513 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-948-564-6

Query Match          10.4%; Score 279.5; DB 3; Length 513;
Best Local Similarity 22.9%; Pred. No. 2.1e-19;
Matches 118; Conservative 82; Mismatches 216; Indels 99; Gaps 17;

QY 15 LPSLLSLFLILKRRNRKTRFNLPPKSGMPFLGETIGYL---KRYTATLGDPMQOH 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15 LAFIFSGILF--FLKOKSKSKKFLPPGPPGMPIVGNLFQVARSGRKFF-----EYNDV 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 72 VSKYGIKIRSNLFEPTIVSADAGLNRFILNLSGRLECSIPR----- 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 68 RLKYSIFTKMGTRTMIILDAKLVHEAMIQGATYATRPENPTRTISENKFYNA 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 115 SIGGILGKWSMLVVGDMHNRMSISLNFSLHARLRTILKDYERHTLFLV-----DSW 168
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 128 TYGPV-----KWSL-----RRNQNMLSSTRLKE--FRSVADNMADKILNKLDEAE 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 169 QONSIFSAODEBAKFTTNLMAKHIMSMDPGEETEOLKKEYVTFMKGVSA----- 219
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 174 KNNGVVWVLDAFRAVCCIVAMCFLEMEDETERIDQ---VMSGLVLTLDPRIDYL 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 PLNLP--GTAVHKALOSRAATILKFTERKMEERKLDIKEEDQEEVEVTEDEAEMSKSDHV 277
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 230 PILSPFSQKQKALEVRREVEFLVPIEORRAIQNPG-----SDHT 273
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 278 RKQRTDDL-----GVLKHSNLSSTEOIIDLTLSLFAGHETSSVAIALAIFLQACP 331
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 ATTFSYLDTLFDLKEVG---KKSAPSDAELVSLCSEFLNGSTDTTAAVEMGIQOLLANP 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 332 KAVEELREEHLEIARAKKEGSELNMDYKKMDFTQCVINETLRAGNVVREFLHRAKALKD 391
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 NVQTKLYEE-----IKRTYGEKKVDEKDKEMPYLHAAVVKELLRKRPPTHFVLTNAVTE 384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 392 -VRKYGYDIPSGMKVLPVTSAVHLDSRYDQPNLFNPMWMOONNGASSSGSFSFTWGN 450
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 PTLIGGTIDIPIDANVEYTPALADPKNNMLNPEKFEDEPRTISGSEADITG-----VTGV 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 451 NMPFGGPRLCAGSELAKLEMAFIHHLVLEKMW 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 KMPFGYGRICPEGLAMATVHIIHIMARMVQOEFRM 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Query Match	9.8%	Score 264	DB 2	Length 520
Best Local Similarity	20.3%	Pred. No. 7.6e-18		
Matches 109	Conservative 103	Mismatches 208	Indels 118	Gaps 18

RESULT 8  
US-08-457-274A-25

Query Match	9.8%	Score 263.5	DB 1	Length 504
Best Local Similarity	24.1%	Pred. No. 8.1e-18		
Matches 132	Conservative 89	Mismatches 213	Indels 113	Gaps 25

275 DHVRKQRTDDDLGWL-----KHSNLTSEQIILDLISLLFAGHETSSVAIALAI 324

```

Db 262 DSVQKHRV--DFQLMMAHNDKSKDESHETALSDMEITAOOSIIFIFAGYEPTSSITLSFVL 319
QY 325 FFLQACPRVAEELREHELEIARAKKEGSESLMNDYKMKDFTQCVINETRLGNAVFL 384
Db 320 HSLATHPTQKKLOE---EIDRALP--NKAPPTDVTWEMEXYLDVNETRLRLPIGRNL 374
QY 385 HRKALKDVRKGYDIPSGWKVLPVISAHLNDSRYDQNLFPNPMWMOOONNGASSSGSGS 444
Db 375 ERVCKKDYEVGFMPKGSVYMIPSYALHRDPQHWPPEEPERPERFSKENGSDIP---- 430
QY 445 FSTGNNNTMPFGGPRILCAGSELAKLEMAVFIHHLVLFKNMELAEDDQPFAPFVDFPENG 504
Db 431 -----YVLPFGNGPRNCIGRMFALMNMKALTLVQLQNFSPQCKETO----- 473
QY 505 LPIRVSR 511
Db 474 IPLKISR 480

```

```

RESULT 9
PCT-US95-05758-25
: Sequence 25, Application PC/TUS9505758
: GENERAL INFORMATION:
: APPLICANT: Cornell Research Foundation, Inc.
: TITLE OF INVENTION: Cytochrome P4501pr Gene and Its
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Nixon, Hargrave, Devans & Doyle
: STREET: P.O. Box 1051, Clinton Square
: CITY: Rochester
: STATE: New York
: COUNTRY: USA
: ZIP: 14603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/05758
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldman, Michael L.
: REGISTRATION NUMBER: 30,727
: REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 716-263-1304
: TELEFAX: 716-263-1600
: INFORMATION FOR SEQ ID NO: 25:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 504 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Rat
: STRAIN: Unknown
: DEVELOPMENTAL STAGE: Adult
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: Unknown
PCT-US95-05758-25

```

Query Match 9.8%; Score 263.5; DB 5; Length 504;  
 Best Local Similarity 24.1%; Pred. No. 8.1e-18;  
 Matches 132; Conservative 89; Mismatches 213; Indels 113; Gaps 25;

```

QY 9 LLLPULLPSLISLLFLILKRRNKRTR--FNLP--PGSKGMPFLGEGTIGYLYKPYTATTL 64
Db 3 LLSALTEITWLLAVLVLLVGLGFTRTGLRPGKQIIPGPKPLPFGVTLNLY-----MGL 57
QY 65 GDFMOQHSKSKGKIRSLNF--GE--PTIYSADAGLNRFLOWEGRLFEC-----STPRSIG- 117
Db 58 WKFDVECHKRYGKITW--GLFDGOMPLFAITTEMLKNVL-----VKECFSVFTNRDGP 110
QY 118 -GILGKMSLVLVGDMHDMRSISLNFSLHARLFTILKDVERRHTLFFVLDSMOONS----- 172
Db 111 VGIMCKAIVSAKDEEMKRYRLLSTFTS--GRKE-MPPIIEQYGDILVKKLQEALENGK 168
QY 173 -----IFSA--ODEAKFTENLMAKHIMS--MDPGEETEQLK--EYVTEMGVVSAPL 221
Db 169 PVTMKKVGAYSMDVITSTSFQVNDVSLNPKDPVEVETKTLRLRDFDPPLFSVLVLPF 228
QY 222 NLPGTAHYKAL-----QSRATILKFTERKMEERKLDIKEDDEEVEVKTEDEAKMSK 274
Db 229 LTP--IYEMLNICMFPRDSIEFKKFVYR--MKETRL----- 261
QY 275 DHVRKORTDDLLGWL-----KHSNLSTEQIIDLILSLFAGHETSSVAIALAI 324
Db 262 DSVQKHRV--DFQLMMAHNDKSKDESHETALSDMEITAOOSIIFIFAGYEPTSSITLSFVL 319
QY 325 FFLQACPRVAEELREHELEIARAKKEGSESLMNDYKMKDFTQCVINETRLGNAVFL 384
Db 320 HSLATHPTQKKLOE---EIDRALP--NKAPPTDVTWEMEXYLDVNETRLRLPIGRNL 374
QY 385 HRKALKDVRKGYDIPSGWKVLPVISAHLNDSRYDQNLFPNPMWMOOONNGASSSGSGS 444
Db 375 ERVCKKDYEVGFMPKGSVYMIPSYALHRDPQHWPPEEPERPERFSKENGSDIP---- 430
QY 445 FSTGNNNTMPFGGPRILCAGSELAKLEMAVFIHHLVLFKNMELAEDDQPFAPFVDFPENG 504
Db 431 -----YVLPFGNGPRNCIGRMFALMNMKALTLVQLQNFSPQCKETO----- 473
QY 505 LPIRVSR 511
Db 474 IPLKISR 480

```

```

RESULT 10
US-08-313-075A-30
: Sequence 30, Application US/08313075A
: Patent No. 5639670
: GENERAL INFORMATION:
: APPLICANT: Holton, Timothy A.
: APPLICANT: Cornish, Regina C.
: APPLICANT: Tanaka, Yoshikazu
: TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Scully, Scott, Murphy & Presser
: STREET: 400 Garden City plaza
: CITY: Garden City
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 11530
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/313,075A
: FILING DATE: 30-NOV-1994
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PL 1538/92
: FILING DATE: 27-MAR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PL 6698/93

```

```

: FILING DATE: 07-JAN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PCT/AU93/00127
: FILING DATE: 25-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Digilio, Frank S.
: REGISTRATION NUMBER: 31,346
: REFERENCE/DOCKET NUMBER: 9433
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 742-4343
: TELEFAX: (516) 742-4366
: TELETYPE: 230 901 SANS UR
: INFORMATION FOR SEQ ID NO: 30:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 476 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-313-075A-30

```

Query Match 9.6%; Score 257; DB 1; Length 476;

Best Local Similarity 24.0%; Pred. No. 3.3e-17; Matches 123; Conservative 87; Mismatches 227; Indels 76; Gaps 21;

```

OY 12 LLLPLSLILLFLILKRRNRKTRFNLPGKSGMPLGETIGYKPYATATGDFMQQH 71
DB 9 LLIYFSLISLKKK-----KKSNCOTK-KLPQGPWKVYFGLSLHMGVGLPHVYLRDL---- 59
OY 72 VSKYKTYRSMLEPPTIVSADAGINRPIIQNEGRLFECSY-PRSIGILKWSMLVLY- 129
DB 60 AKKYGPIHMLQGLKISAVVTSPEAKRKVLKTHDLAF-AVRPKLLGLEIYCYNSSDIAP 117
OY 130 ---GDMHDMASIS-LNFLSHARLRTILKDYERHTLVLDWMQNS-----IFS 175
DB 118 SPYGDYWRMKRICYLEVLSAKNWS--FNSIRDEILIMIDFLRSSGKPYNITERIFS 175
OY 176 AQDEAKRTFIMAKHIMSMPGE-EETEOALKREYTFMKGVSAFLNPGTAY-HKALQ 233
DB 176 -----FTSSMICRSVGRKRIKEDCIRHKYKKTGLIDGVDADT-FPSIRFLHVLIG 227
OY 234 SRATLTKIERKMEERKLDIKEEDQEEBEVKTDEBAENKSDHVRKQRTDLDLGMVLKH 293
DB 228 MKGRIMD-VHKKV-----DAIVEYNNNEHK-ETLRGTGTNGEYGGEDLDIVLRL 275
OY 294 SN-----LSTEOILDLILSLPAGHTSSVAIALAIFFLQACKAVEELREHLEIAR 346
DB 276 KEEDGLQPLPINDIKALFNDMFAGTSTTTIMMAVELMKNPVSFAKAQAEVREYVK 335
OY 347 AKKELGESELMWDYKKKDFQOCVINETLRL-GNVVFLHRRALKDVRKYGDIPIGMYK 405
DB 336 GKPEFDE-----DDEELNLYKLVIREFLRHLPLPLPRECRRETELNGTITPLANKV 390
OY 406 LPVISAVALHDSRYDQPLFNPWRMOQONNGASSGSGFTSWGNN--YMPFGGSPRLCA 463
DB 391 IVNWALIGRDKRYMDASFKPERPEH-----NSLNFAGNFEYELPFGSSRRICP 440
OY 464 GSELAKLEMAVFIHHLVLYKFWNELAEDDQPAF 496
DB 441 GISFGLANVYHPLAQLVYHFMRLPTGVDPNDF 473

```

```

: RESULT 11
: US-08-201-118-7
: Sequence 7, Application US/08201118
: Patent No. 5786191
: GENERAL INFORMATION:
: APPLICANT: GOLDSTEIN, Joyce A.
: APPLICANT: ROMKES-SPARKS, Marjorie
: TITLE OF INVENTION: CLONING AND EXPRESSION OF COMPLEMENTARY
: TITLE OF INVENTION: DNAS FOR MULTIPLE MEMBERS OF THE HUMAN CYTOCHROME P450 2C
: NUMBER OF SEQUENCES: 44

```

```

: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Kourie and Crew
: STREET: 379 Lytton Avenue
: CITY: Palo Alto
: STATE: California
: COUNTRY: US
: ZIP: 94301
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/201,118
: FILING DATE: 22-FEB-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/864,962
: FILING DATE: 09-APR-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Liebeschuetz, Joe
: REGISTRATION NUMBER: 37,505
: REFERENCE/DOCKET NUMBER: 15280-192-1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2422
: TELEFAX: (415) 326-2400
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 490 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-201-118-7

```

Query Match 9.6%; Score 256.5; DB 1; Length 490;

Best Local Similarity 20.8%; Pred. No. 3.9e-17; Matches 110; Conservative 99; Mismatches 207; Indels 119; Gaps 20;

```

OY 11 PLLPLSLILLFLILKRRNRKTRFNLPGKSGMPLGETI-----GYKPYATATLG 65
DB 3 FEVVLVLCLSFMLFSLMRQSCRKRL-LPQPTPLPIGNMLQIDVQDICKSF----- 55
OY 66 DEMQOHVSK-YGKITRSMLEPPTIVSADAGINRPIIQNEGRLFECSYPRSIGILK-W 123
DB 56 -----NFSKYVGPVPTVYFGMPDIVFHYGEAVKALIDNGEEFGSGNSPIQRITKGL 110
OY 124 SMLVLYGDMHDMRISINFLSHARLRTILKQD---VERHTL-----FVL 165
DB 111 GIISNGKRWKEIRFSLTNLRNFGMGKRSIEDRYQOEAHCLVEELRKTKASPCDPTFL 170
OY 166 DSMQNSIFSADAEAKKTFEFLMAKHIMSMPGEBETQLKREYVTFMK-----GVVSA 219
DB 171 GCAPCNVICSVYFQ-KREDYK-----DONFLYLMKRFENENRIINS 210
OY 220 PL-----NLPGTAYHKAQOSRATILKTERKMEERKLDIKEEDQEE-----V 263
DB 211 PWIOVCNNEPRLIDCFPGT-HNKVLKNAVALRYSYIREVKEHQASLDVNNPDEFMDCEFLI 269
OY 264 KTEDEAENKSDHVRKQRTDLDLGMVLKHSNLSTEOILDLILSLPAGHTSSVAIALA 323
DB 270 KMEQEKDNKSF-----FNLENLVGYADLFVAGTETSTTLIRG 309
OY 324 IFFLQACKRAVEELREE--HEIARAKKELGESELMWDYKKKDFQOCVINETLRLGNVY 381
DB 310 LLLLLKHEVTAQVOEEDIHV-IGRHSRSC-----MDRSHPYTDVAVHEIQYSDLY 362
OY 362 -RFLHRRKALKDVRKYGDIPIGMYKVLPIYSAVHLNDSRYDQPLFNPWRMOQONNGASSS 440
DB 363 PTGVHAATVTDKRFKNVLYPKGTTIMALLTSYLDHDKPEPNINIPDPGHFDLKN----- 416
OY 441 GSGSFTWGNMMPFGGSPRLCAGSELAKLEMAVFIHHLVLYKFWNELAED 490

```

DB 417 -GNFK-SDYFMPFSAGKRICAGEGLARMLEFLTITLQNFNLKSVSD 463

RESULT 12

US-08-238-821B-7

Sequence 7, Application US/08238821B

Patent No. 5912120

GENERAL INFORMATION:

APPLICANT: GOLDSTEIN, Joyce A.

APPLICANT: ROMKES-SPARKS, Marjorie

APPLICANT: DE MORALS, Sonia M.F.

TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN

TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT OF S-

NUMBER OF SEQUENCES: 61

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th floor

CITY: San Francisco

STATE: California

COUNTRY: US

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/238,821B

FILING DATE: 06-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/201,118

FILING DATE: 22-FEB-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/864,962

FILING DATE: 09-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Liebeschuetz, Joe

REGISTRATION NUMBER: 37,505

REFERENCE/DOCKET NUMBER: 15280-192110US

TELEPHONE: (650) 326-2400

TELEFAX: (650) 326-2422

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 490 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-238-821B-7

Query Match 9.6%; Score 256.5; DB 2; Length 490;

Best Local Similarity 20.8%; Pred. No. 3.9e-17;

Matches 110; Conservative 99; Mismatches 202; Indels 119; Gaps 20;

DB 11 PLULLPSLLFLFLILKRRNRKTRNLPKSGMFLGFTI-----GLKRYTATTLG 65

DB 3 FVAVLYCLTSFMLEFSLIMROSCRRK--LPPGPPPLPTIGNMLQIDVKICKSFT----- 55

DB 66 DFMQOHVSK-YGRIYRSNLGEPTIYSADAGLNRFLQNGRGLFECSYPSISGILGK-W 123

DB 56 -----NSKYVGPVETFYFPGMPLVFGHEAVKKAALIDNGEESGSGNSPIQORITKGL 110

DB 124 SMALVYGDMDRMRISLNLFLSHARLTITLKD---VERHTL-----EVL 165

DB 111 GIISNCKRKKEIRRFSLTILRNFGMKRSIEDRVOEAEHCLVEELKTKRASPDPFFIL 170

DB 166 DSMQONSIFSAOEAKKFTFNLAHAKHMSMDPBEETEOJLKEYVTFMK-----GVVSA 219

DB 171 GCAPCNVCISVTEQ-KRFDPK-----DQNFLLMKRFENFRILNS 210

DB 220 PL-----NLPGTAVHKAJOSATILKFERKMEERKLDIKEEQEERE-----V 263

DB 211 PMIOVCNNPFLILDCFEFGT-HNKYLKNAVLTIRSTIRKRYVEHQASLDVNNPRDPMGFLI 269

DB 264 KTEDEAEMSKSDHYRKQRTDDLLGVNLKHSNLSTEOILDLISLAFAGHETSVAIALA 323

DB 270 KMEQKDNQKSE-----FNIEMLVGTADVLEFVAGTETSTTLRYG 309

DB 324 IFPLQAPKAVEELRER-LEIETAKKELGSESLMNDVKKKDFQCVINETLRGNV 381

DB 310 LLLLKHPEVTAKVQEEIDIV-IGRRSPC-----MQRSHMPTIDAVVHEIQORSDLV 362

DB 382 -RFLHRKALDVRKYGYDIPSGWKVLPVISAHLDNSRDQPLNFMWMOOONCASS 440

DB 363 PTGVPNAVTTDTKFRNRLIRKGTITMALNLSVLDHDEFFNPFIIDFGHLDKN----- 416

DB 441 GSGSFTWGNMNYMFGGGPRLCAGSELAKLEMAVEIHHLVLFKNWEIAED 490

DB 417 -GNFK-SDYFMPFSAGKRICAGEGLARMLEFLTITLQNFNLKSVSD 463

RESULT 13

PCT-US95-05744-7

Sequence 7, Application PC/TUS9505744

GENERAL INFORMATION:

APPLICANT: GOLDSTEIN, Joyce A.

APPLICANT: ROMKES-SPARKS, Marjorie

APPLICANT: DE MORALS, Sonia M.F.

TITLE OF INVENTION: CLONING, EXPRESSION AND DIAGNOSIS OF HUMAN

TITLE OF INVENTION: CYTOCHROME P450 2C19: THE PRINCIPAL DETERMINANT

NUMBER OF SEQUENCES: 61

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Crew

STREET: 379 Lytton Avenue

CITY: Palo Alto

STATE: California

COUNTRY: US

ZIP: 94301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/05744

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/238,821

FILING DATE: 06-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/201,118

FILING DATE: 22-FEB-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/864,962

FILING DATE: 09-APR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Dow, Karen B.

REGISTRATION NUMBER: 29,684

REFERENCE/DOCKET NUMBER: 15280-192-1-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 326-2400

TELEFAX: (415) 326-2422

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 490 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein  
PCT-US95-05744-7

Query Match	9.6%;	Score 256.5;	DB 5;	Length 490;
Best Local Similarity	20.8%;	Pred. No. 3.9e-17;		
Matches 110; Conservative	99;	Mismatches 202;	Indels 119;	Gaps 20;

[illegible]

RESULT 14  
 US-08-532-065B-2  
 : Sequence 2, Application US/08532065B  
 : Patent No. 5753507  
 : GENERAL INFORMATION:  
 : APPLICANT: Ohta, Daisaku  
 : APPLICANT: Mizutani, Masaharu  
 : TITLE OF INVENTION: Plant Geraniol/Nerol 10-hydroxylase and  
 : TITLE OF INVENTION: DNA Coding Therefor  
 : NUMBER OF SEQUENCES: 6  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: No. 5753507artls Corporation  
 : STREET: 59 Route 10  
 : CITY: East Hanover  
 : STATE: NJ  
 : COUNTRY: USA  
 : ZIP: 07936  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: IBM PC compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/532,065B  
 : FILING DATE: 22-SEP-1995  
 : CLASSIFICATION: 800  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Meigs, J. Timothy

```

? REGISTRATION NUMBER: 38,241
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 919-541-8587
? TELEFAX: 919-541-8669
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 495 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-532-065B-2

```

Query Match	9.1%;	Score 245;	DB 1;	Length 495;
Best Local Similarity	22.9%;	Pred. No. 5.6e-16;		
Matches 120;	Conservative 80;	Mismatches 218;	Indels 106;	Gaps 19;

```

QY 12 LLLEPSSL-LLLELILIKRNRKTRPNLPKSGMPELITIGYLKPYATTLGDMQ 70
D 10 LLLEPSSL-LLLELILIKRNRKTRPNLPKSGMPELITIGYLKPYATTLGDMQ 67
QY 71 HVSKYKTYKRSNLFGEPTIYSADGLNFIQNGRLEFESYP---RSIG----- 117
D 68 ---TYGPMYSUKLSLNTVVIVIASPEAREVYRTQDILSARSPNAVRSINHODASLWLT 124
QY 118 -GILGKMSMLVLYGDMHRMDSISL-NPLSHARL---RTILKIVERTHTFVLDSMOONS 172
D 125 PSSSARMLL-----RLSYTQLLSQRLEAKRKALRMNVKELVFISSSDREE 174
QY 173 IFSADKAKFTFNLMAKNHMSMDPEEBEETOLKEVYTKGVASAPLPGT----- 226
D 175 SVDISRVAFITTLNISILFESVDLSGYN---KASINGVODIVISVMAAGTPDANY 230
QY 227 ---AVHKAOSRATILK-----FIERKMEERKLIDKEQDDEEBVKEDBAEM 271
D 231 FPLFLFDLQGNVTKFYKCTELRYVRGFIIDAKIAESSONPKD-----V 277
QY 272 SKSDHVRKORDDDLGLGWLKHSNLSTEOULDLILSLFAGHETSVAVALAIFLOACP 331
D 278 SKNPFV-----DNULDYKGESELSIDIEHLIDMTACTDTSSTLEWPMTELLKNP 331
QY 332 KAVBELREHLEIARAKE---LGESL-NMDDYKKMDFQCYINFTLRIGNVYRL-H 385
D 332 KT-----MAKQAEIDCVIGQNGIYVESDISKILPYQAVVAKKEFFRHTHPVLLIP 381
QY 386 RKAALDYKYGIDYIPSGKAGLPVJISAVHLNDRSARDOPMLFPMYMOOONNGASSGSGSF 445
D 382 RKAESDAIILGLFWLKDQVLVYVMAIGRDPDSVMDNPQFEPERFLGD-----M 433
QY 446 STWGNV---MPFGGPRICASASELAKEMAVFIHHLVYKFMWEL 487
D 432 DVRGDVELTFFGAGRRICCPMPILAMKTYSIMLASLVSFWML 475

```

RESULT 15  
 US-08-948-564-16  
 : Sequence 16, Application US/08948564  
 : Patent No. 6121512  
 : GENERAL INFORMATION:  
 : APPLICANT: Siminszky, Balazs  
 : APPLICANT: Dewey, Ralph E.  
 : APPLICANT: Corbin, Frederick T.  
 : TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and  
 : TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants  
 : NUMBER OF SEQUENCES: 23  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Virginia C. Bennett  
 : STREET: PO Box 37428  
 : CITY: Raleigh  
 : STATE: No. 6121512th Carolina  
 : COUNTRY: USA  
 : ZIP: 27627  
 : COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/948,564  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Virginia C.  
REGISTRATION NUMBER: 37,092  
REFERENCE/DOCKET NUMBER: 5051-409  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-854-1400  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 576 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-948-564-16

Query Match 9.18; Score 244.5; DB 3; Length 576;  
Best Local Similarity 22.7%; Pred. No. 7.9e-16;  
Matches 111; Conservative 91; Mismatches 181; Indels 107; Gaps 22;

QY 63 TLGDFMOQHYSKYGIYRSLNFGPTIVSADAGLNRFILQ-----NEGRLEFEGSYPRSI 116  
DB 97 SLIDWFELEH-----GAVYKLFPGKAFYVSDPIYARHILRENAFSYDKGLADILEP--- 149  
QY 117 GILGK-----WSM--LYLVGDMHRMRSISLNLSHARLITLL--KDYERHTLF 163  
DB 150 --IMGKGLIPADLTQWQRRRVIAPAFHNSYLEAMVKIFTTCSERTILKFNKLLLEGGYD 207  
QY 164 VLDSMOQNSIFSAODEAKKFTFNMAKHISMOPGE--EETEOUKREYVFMKGVASAPL 221  
DB 208 GPDSIE---LDLEAFSSSLALDIIIGLVFNYPDGSVTKESPVIKAVYGTLFEAHERSTF 263  
QY 222 NLPGTAYHKAQSRATILKFTERKMEERKLDIK-----EDQEEEEVKTEDE 268  
DB 264 YIP---YWKIPLARWVYPR--QKRFPD--DLKYINTCLDGLIRNAKESROETDVE---- 311  
QY 269 AEMSKSDHVRKQRTDDLLGMV--LKHNSLSTEQIIDLILSLFAGHETSSVAIALAIF 326  
DB 312 -KLQQRDYLNLIK--DASLLRFLVDMRGADVDRLRDLMTMLAGHETTAAYLTWAVFL 368  
QY 327 LQACPKAVEELREHLEIARAKKELGESELMWDYKKMDFTQCVINETLRIGNVRFILHR 386  
DB 369 LAONPSKMKRAQ-----AEVDLVLTGTRPFESLKELOYIRLIVEALRLYPOPLLIR 422  
QY 387 KALK-DY-----RYKGYDIPSGWKVLPYISAVHLDSRY--DQPNLFNPMWQOONG 436  
DB 423 RSLKSDVLPGGHKEGKGYALPAGTDVF--ISYVNLHRSPLYFMDRPPDFEERLYQNN 480  
QY 437 ASSSGSGSESTWGN-----NYMPFGGPRICAGSELAKLEMAVFIH 477  
DB 481 EEIEG-----WAGLDPSRSPGALYPNEVISDFAFLPFGGPRKCVGDQFALMESTVALT 534  
QY 478 HLYLKFNNMEL 487  
DB 535 MLQONFDVEL 544

Search completed: September 28, 2001, 18:44:39  
Job time: 69 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2001, 18:43:30 ; Search time 22.2 Seconds  
(without alignments)  
1400.906 Million cell updates/sec

Title: US-09-502-426-2

Perfect score: 2681  
Sequence: 1 MFETHTHTLLPLLLPSLLS.....FAFPVDFPNGPLRYSRLL 513

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_0601.\*  
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT.\*  
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT.\*  
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT.\*  
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT.\*  
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT.\*  
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT.\*  
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT.\*  
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT.\*  
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT.\*  
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT.\*  
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT.\*  
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT.\*  
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT.\*  
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT.\*  
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT.\*  
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT.\*  
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT.\*  
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT.\*  
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT.\*  
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT.\*  
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT.\*  
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2681	100.0	513	21	AAB07921 A cytochrome P450
2	978.5	36.5	472	18	AAW27153 Arabidopsis thalia
3	978.5	36.5	472	21	AAG44571 Arabidopsis thalia
4	978.5	36.5	472	21	AAG45022 Arabidopsis thalia
5	978.5	36.5	491	21	AAG45021 Arabidopsis thalia
6	978.5	36.5	492	21	AAG44570 Arabidopsis thalia
7	946	35.3	444	21	AAG44572 Arabidopsis thalia
8	946	35.3	444	21	AAG45023 Arabidopsis thalia
9	726	27.1	461	21	AAG46490 Arabidopsis thalia
10	726	27.1	462	21	AAG46489 Arabidopsis thalia
11	726	27.1	465	21	AAG46491 Arabidopsis thalia

12	712	26.6	461	21	AAG11836 Arabidopsis thalia
13	712	26.6	462	21	AAG11835 Arabidopsis thalia
14	712	26.6	465	21	AAG11834 Arabidopsis thalia
15	653.5	24.4	468	21	AAG20783 Arabidopsis thalia
16	606	22.6	471	21	AAG23013 Arabidopsis thalia
17	606	22.6	479	21	AAG23012 Arabidopsis thalia
18	603	22.5	469	21	AAG23014 Arabidopsis thalia
19	598.5	22.3	430	21	AAG20784 Arabidopsis thalia
20	547.5	20.4	489	21	AAG30048 Arabidopsis thalia
21	546	20.4	481	21	AAG30049 Arabidopsis thalia
22	528.5	19.7	388	21	AAG20785 Arabidopsis thalia
23	503	18.8	492	19	AAW37733 Arabidopsis thalia
24	503	18.8	492	19	AAW44159 Arabidopsis thalia
25	490	18.3	433	21	AAG30050 Arabidopsis thalia
26	461	17.2	497	19	AAW37735 Arabidopsis thalia
27	461	17.2	497	19	AAW44161 Arabidopsis thalia
28	439	16.4	497	19	AAW37734 Arabidopsis thalia
29	439	16.4	497	19	AAW44160 Arabidopsis thalia
30	420	15.7	253	21	AAG17903 Arabidopsis thalia
31	411.5	15.3	237	21	AAG17905 Arabidopsis thalia
32	411.5	15.3	242	21	AAG17904 Arabidopsis thalia
33	319	11.9	519	21	AAG42953 Arabidopsis thalia
34	318	11.9	519	21	AAG16752 Arabidopsis thalia
35	316	11.8	500	21	AAG42954 Arabidopsis thalia
36	315	11.7	500	21	AAG16753 Arabidopsis thalia
37	298.5	11.1	905	21	AAG38596 Arabidopsis thalia
38	298.5	11.1	949	21	AAG38595 Arabidopsis thalia
39	298.5	11.1	1041	21	AAG38594 Arabidopsis thalia
40	297	11.1	501	21	AAG50190 Arabidopsis thalia
41	297	11.1	516	21	AAG50189 Arabidopsis thalia
42	290.5	10.8	502	21	AAG29761 Arabidopsis thalia
43	289	10.8	511	20	AAV23341 A P450-2 protein 1
44	289	10.8	511	21	AAV19694 Arabidopsis thalia
45	285.5	10.6	502	21	AAV91348 Human secreted pro

#### ALIGNMENTS

RESULT 1	
AAB07921	
ID AAB07921 standard; Protein: 513 AA.	
XX	
AC AAB07921;	
XX	
DT 14-NOV-2000 (first entry)	
XX	
DE A cytochrome P450 enzyme designated DMF4.	
XX	
KW DMF4: cytochrome P450 enzyme; brassinosteroid; 22alpha-hydroxylation;	
KM plant phenotype; cell elongation.	
XX	
OS Arabidopsis sp.	
XX	
PN WO200047715-A2.	
XX	
PD 17-AUG-2000.	
XX	
PF 11-FEB-2000; 2000WO-US03820.	
XX	
PR 11-FEB-1999; 99US-0119657.	
PR 11-FEB-1999; 99US-0119658.	
XX	
PA (ARIZ-) ARIZONA BOARD OF REGENTS.	
XX	
PI Azpilroz R, Choe S, Feldmann KA;	
XX	
DR WPI: 2000-549142/50.	
DR N-PSDB; AAA59599.	
XX	
PT New isolated dmf4 polynucleotide useful for altering the phenotype of	
PT plants, for diagnostic assays and in the production of antibodies -	
XX	

PS Claim 50: Fig 11; 113pp; English.

XX The present sequence represents a DWf4 polypeptide. The polypeptide is a

CC cytochrome P450 enzyme that mediates multiple steps in synthesis of

CC brassinosteroids. Specifically, it mediates multiple

CC 22alpha-hydroxylation steps in brassinosteroid biosynthesis. The DWf4

CC polynucleotide is used for altering the phenotype of a plant. DWf4

CC plants display a dramatic reduction in the length of different organs,

CC and this size reduction is attributable to a defect in cell elongation.

CC The DWf4 polynucleotides and polypeptides can be used in diagnostic

CC assays and to generate antibodies, which can be used to produce

CC immunogenic compositions.

XX

XX Sequence 513 AA:

Query Match 100.0%; Score 2681; DB 21; Length 513;

Best Local Similarity 100.0%; Pred. No. 6.6e-229;

Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFETHTLLPLLLPLLSLLFLLLKRRNRKTRFNLPPGKSGWPLGTTGYLKPYT 60

DB 1 mfehtllpllllpsllsllflllkrrnrktrfnlppgksgwpllgtylkpyt 60

QY 61 ATTLDGMQOHVSKYKIVSNLGEPTVSADAGLNRFILONGRLEFCSYPSISGIL 120

DB 61 attldgmqohvskykivsnlgeptivsadaglnrflfongrlefcyspsisgil 120

QY 121 GKMSMLVLDMDHMRMSISLNFSLHARLRTILKDYERHTFLVLDGMOONSIFSADDEA 180

DB 121 gkmsmlvldmdhmrmsisl n fslharlrtl lkdyerhtfl vldgmoonsif s addea 180

QY 121 gkwamlyvldmhrmsislnfshartrtl lkdyerhtfl vldgmoonsif s addea 180

DB 181 KKTFTNMAKHMSMDPEETEOLEKKEYVTFMKGVVSAPLNLPGTAYHKALOSRAITLK 240

DB 181 kktftnma khmsmdpeet eolekkeyvtf mkgvvsaplnlpgtayhka losratilk 240

QY 241 FIERKMEERLDIKREOGEFEVTEDEAEKMSKSDHVRKORTDDLLGWVLKHSNLTSEQ 300

DB 241 fierkmeerldikre ogefevte deae kmskshvrkgrtd d l l g w v l k h s n l s t e q 300

QY 301 ILDLISLFLAGHETSSVAIALAIFFLQACPKAVEELREHELETARAKKEGSELNMD 360

DB 301 lldlislflaghetssvaialaifflqacpkaveelreheletarakke gselnmd 360

QY 361 YKKMDFTQCVINETLRIGNVVRFLLHKKALKDVRKYKGYDIPSGWKVLPVISAVHLNDRYD 420

DB 361 ykkmdftqcvinetlrignvvrflhkkalkdvrkykgydipsgwkvlpvisavhln d r y d 420

QY 421 OPNLFNFMWROQONNGSSSGSFSFTWGNVMPFGGPRLLCAGSELAKELEMAVFTHHLY 480

DB 421 opnlfnfmr oqonngsssgsfsftwgnvmpfggprllcagseleakel emavfthhly 480

QY 481 LKFNWELAEDDQFPAPFVDFPNGLPIRVSRIL 513

DB 481 l k f n w e l a e d d q f p a p f v d f p n g l p i r v s r i l 513

RESULT 2

AAW27153 standard; Protein; 472 AA.

XX

XX AAW27153:

XX

DT 14-APR-1998 (first entry)

XX

XX Arabidopsis thaliana cytochrome P450-type hydroxylase.

XX

KM Cytochrome P450-type hydroxylase; identification; brassinosteroid;

KW brassinosteroid inhibitor; modified plant; recombinant production;

XX

XX Arabidopsis thaliana.

XX

PN W09735986-A1.

XX

PD 02-OCT-1997.

XX

PF 27-MAR-1997: 97WO-EP01586.

XX

PR 27-MAR-1996: 96US-0622166.

XX

PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX

PI Altmann T, Koncz C, Mathur J, Szekeres MA;

XX

DR WPI: 1997-489649/45.

DR N-PDB: AAT85306, AAT85307.

XX

PT New isolated plant cytochrome P450-type hydroxylase gene - used to

PT identify substances acting as brassino-steroid(s) or brassinosteroid

PT inhibitors for the production of modified plants

PS Claim 1: Pages 44-46; 77pp; English.

XX

XX The present sequence is Arabidopsis thaliana cytochrome

CC P450-type hydroxylase. The hydroxylase can be used to identify

CC brassinosteroids or brassinosteroid inhibitors, useful to produce

CC plants with modified physiological and/or phenotypic

CC characteristics. The modified plants may show, e.g. stimulation of

CC growth, increased cell elongation, increased wood production,

CC accelerated seed germination at low temperatures, an increase in

CC dry weight, repressed anthocyanin production during growth in light

CC and/or inhibited de-etiolation which is induced, e.g. by cytokinin,

CC in the dark or an increase in stress tolerance. The hydroxylase or

CC its coding sequence can also be used for the recombinant production

CC of compounds, e.g. teasterone.

XX

XX Sequence 472 AA:

Query Match 36.5%; Score 978.5; DB 18; Length 472;

Best Local Similarity 41.9%; Pred. No. 3e-78;

Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

QY 12 LLLPLSLLLPLFLILKRRNRKTRFNLPPGKSGWPLGTTGYLKRYTATLGDPMQOH 71

DB 7 lllpsllsllplflil krrnrktrfnlppgksgwpllgtylkrytatlgdpmqoh 71

QY 72 VSKYKIVRSNLFGEPTVSADAGLNRFILONGRLEFCSYPSISGILGKWSMLVLVD 131

DB 72 vskykivrsnlfgeptivs adaglnrfl ongrlefcyspsisgilgkwsmlvlvd 131

QY 64 varygsvfmhlfgeptisadepetnrfvignegklfecspasincnlghshllmkgs 123

DB 64 varygsvfmhlfgeptisadepetnrfvignegklfecspasincnlghshllmkgs 123

QY 132 MHRDMSISLNFSLHARLRTILKDYERHTFLVLDGMOONSIFSADDEAKKFTFNMAKH 191

DB 132 mhrdmsisl n fslharlrtl lkdyerhtfl vldgmoonsif s addea k k f t f n m a k h 191

QY 124 LKTRHSLTMSFANSILKHLMDIDRLVRFINDSWSSRVLL--MEAAKKITFELVKQ 181

DB 124 lktrhsltmsf ansil khlmdidrlvrf indswssrvll -meaakkitfelvkq 181

QY 192 IMSMDPEETEOLEKKEYVTFMKGVVSAPLNLPGTAYHKALOSRAITLKFERKMEERKL 251

DB 192 imsm dpeet eolekkeyvtf mkgvvsaplnlpgtayhka losratilkferkmeerkl 251

QY 182 lmsfdpg-ewsealrkeyllviegffslpplfstykykaiga-----tkvaealt 232

DB 182 lmsfdpg-ewsealrkeyllviegffslpplfstykykaiga-----tkvaealt 232

QY 252 DIRKEDQEEFEVTEDEAEKMSKSDHVRKORTDDLLGWVLKHSNLTSEQILDLISLFLA 311

DB 252 dirkedqee fevte deae kmskshvrkgrtd d l l g w v l k h s n l s t e q i l d l i s l f l a 311

QY 233 vvmktrreeeegae-----tkkdmllaalaaaddgfsadeelvdffvallva 278

DB 233 vvmktrreeeegae-----tkkdmllaalaaaddgfsadeelvdffvallva 278

QY 312 GHETSSVAIALAIFFLQACPKAVEELREHELETARAKKEGSELNMDYTKMDFTQCVI 371

DB 312 ghets svaialaifflqacpkaveelreheletarakke gselnmdytkmdftqcv i 371

QY 279 gyetstlmflavkflteplalaqkkeeklramksd--syslewsdykmpftqcv 336

DB 279 gyetstlmflavkflteplalaqkkeeklramksd--syslewsdykmpftqcv 336

QY 372 NETLRIGNVVRFLLHKKALKDVRKYKGYDIPSGWKVLPVISAVHLNDRYDQPNLFNPMRWQ 431

DB 372 netlrignvvrflhkkalkdvrkykgydipsgwkvlpvisavhln d r y d q p n l f n p m r w q 431

QY 337 netlvraanllygyvrrtamctvelkyklygwkvfssfrvnhlompnkfkarctfnprwq 396

DB 337 netlvraanllygyvrrtamctvelkyklygwkvfssfrvnhlompnkfkarctfnprwq 396

QY 432 QONNGASSSGSFSFTWGNVMPFGGPRLLCAGSELAKELEMAVFTHHLYLKFNMWELAE 490

DB 432 qonngasssgsfsftwgnvmpfggprllcagseleakel emavfthhlyl k f n m w e l a e d 490

QY 397 -----snsvtgpnavvfcpfggprllcpgyelaralavflhrlvtglsawypaeq 446

DB 397 -----snsvtgpnavvfcpfggprllcpgyelaralavflhrlvtglsawypaeq 446

OY 491 DQFAFPVDFPNCGLPIRYSR 511  
| : | | | |  
Db 447 dklvffpltrtqtkryplfvkr 467

RESULT 3  
ID AAG44571  
AAG44571 standard; Protein: 472 AA.

AC AAG44571;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 55847.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
P9 EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PE 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.

```
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 36.5%; Score 978.5; DB 21; length 472;
Best Local Similarity 41.9%; Pred. No. 3e-78;
Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;
```

```
OY 12 LLLPSSLILFLILKRRNRKTRFNLPRKSGWPLFGFTIGVLRKYTRTTLGDEFQOH 71
DB 7 llllssaaafll--lllrrtyrmgllppssjgllpllgelfqligaaykxtenpeptl 63
OY 72 VSKRYGKIYRNLGCEPTIVSADAGLNRFLIQNGRGLFECSYPRSIGLIGKWSMLVGD 131
DB 64 varygsfwmhllgpeplfsadpetnrfvlnqnegklfecyspasicnllgkhslllmkgs 123
```

```
OY 132 MHRMRSISLNFSLHARLRTLLKDYERHTLFVLDSSQONSISFADPAKFFENLMAKH 191
DB 124 lhrmrhslmsfnsnsllkohlmdldtrlyrfdssvsvll--meaakkltfelvkq 181
OY 192 IMSMDPEEETEDOLKRYVTFPMKGVSAAPLNLPGTAHKLQSRATLTKFERKMEERKL 251
DB 182 lmsfdpg-ewseslrkeyllviegelfsdlprlflsfatykrkaqar-----fkvaeatl 232
OY 252 DIKEEDOEEREVTEDAEKSKSDHVAKKORTDDLLGWLKHSLSLSTEOILLDLSLFA 311
DB 233 vvmkrrreeegae-----fkkmllaalaaaddgfsdeeltvdlvalva 278
OY 312 GHETSSVAIALAIFFLQACPKAVEELREEHLEIARAKKEELMDDYRKMDFTQCVI 371
DB 279 gyetstlmclavkflletpqlaqlkeehkrlamksd--syjlewdsyksmpftgcgv 336
OY 372 NEPLRLGNVYRFLHRKALKVRYKGYDIPSGWKVLPVSAVHLDNSRYDDPNLFPWRMQ 431
DB 337 netlrvanllgvyfrlramtvelkylkprgwkvyfssfravhldpnfkdarlfnprwq 396
OY 432 QONNGASSGSGSFSTWGN--YMPFGGPRLCAGSELAKEAMVFIHVLKFNWELAED 490
DB 397 -----snsvltgprsnvltprfg99prlcpgyelarvalsvflnlvtgfswpaeq 446
OY 491 DQPFAPFPVDFPNGLPIRVSR 511
DB 447 dkilvffptltqkrpflvkr 467

RESULT 4
AAG45022
ID AAG45022 standard; Protein; 472 AA.
XX
AC AAG45022:
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 56469.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132047.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
```

PR	07-MAY-1999;	99US-0132863.
PR	11-MAY-1999;	99US-0134256.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134219.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139763.
PR	22-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140891.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	26-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145951.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.

PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Query Match

36.5%; Score 978.5; DB 21; Length 472;

Best Local Similarity 41.9%; Pred. No. 3e-78;

Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;

OY 12 LLLPPLSLLEFLILKRRNRKTRNLPBGKSGWPLGEGTIGLKPYATATLIGDMQOH 71  
| | | | : | | | | | | | | | | | | | | : |  
DB 7 lllslaaqfl---lllrrtyrmglppgsiglligetflligayktenpepfider 63  
| | | | : | | | | | | | | | | | | | | : |  
OY 72 VSKYGRIVSNLGEPEPTIVSADAGLNRFILQNGRLFECSYPRSIGIGLIGKSMVLVGD 131  
| : | | : : : | | | | | | | | | | | | | | : |  
DB 64 varygsvmfchltgepctlsadepetrfrvlgneqkllfecsypasichlllgkshlllmks 123  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 132 MHRDMRSISLNLFSHARLRTILKDYERHTLEVLDSMQNSIFSADDAKKEFTNMAKH 191  
| : | | : : : | | | | | | | | | | | | | | : |  
DB 124 lhrmrlmsfmsfnssllkdhlmldldrvfrldswssvll--meakkltfelfvkg 181  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 192 INSMDDGEETEOLEKKEYVTFMKGVVSAPLNDGTAYHKAQSRATILKFERKMEERKL 251  
| : | | | | : : | | | | : : | | | | | | | | : |  
DB 182 lmsfdpg-ewseslrkeyllvlegfslpplfstctyrkaidqr-----rkvaalt 232  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 252 DIKEEOEVEEVTEDEAEKMSKSDHVRKORTDDLLGWVLKHSNLSPEOILLILSLFLPA 311  
| : | | : : : | | | | | | | | | | | | | | : |  
DB 233 vvmkrrreeeeagae-----rkkmlaallaaadgfdedeelvalialva 278  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 312 GHETSSVAIALAIFLQACPKAVEELREHLEIARAKKEGESEELNWDYDKKMDFTQCVI 371  
| : | | : : : | | | | | | | | | | | | | | : |  
DB 279 gyetctatmlavkflteplrlaqkleehekltamskd--syslwsdykxmpfcgv 336  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 372 NEFLRGVNVRFHRAKALMDVRKYGDIPSGKWVLVVISVHNDNSRYDDPNLFNWRMQ 431  
| | | | : : : | | | | | | | | | | | | | | : |  
DB 337 necllvaanllgyvframcdveikykllpgwkvfssfravhldphfkdaftfmrwg 396  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 432 OQNNGASSSGSFSWGN--YMPFGGPRILCAGSEIATLEMAVFTLHILVLEKFNWELAD 490  
| : | | : : : | | | | | | | | | | | | | | : |  
DB 397 -----snsvtgpnavvfcpfggprlcpgyelarvalsvflhrlvtgfswpag 446  
| : | | : : : | | | | | | | | | | | | | | : |  
OY 491 DQPFAPFVDFPNGLPRIYSR 511  
| : | | : : : | | | | | | | | | | | | | | : |  
DB 447 dkivffltrtkryplfvkr 467  
| : | | : : : | | | | | | | | | | | | | | : |

## RESULT 5

AAC45021 standard; Protein; 491 AA.

AAC45021;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 56468.

Protein identification: signal transduction pathway: metabolic pathway:

hybridisation assay; genetic mapping; gene expression control; promoter;

Arabidopsis thaliana.

EPI033405-A2.

XX

PD 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 23-APR-1999; 99US-0130891.  
XX 28-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 04-MAY-1999; 99US-0132407.  
XX 04-MAY-1999; 99US-0132484.  
XX 05-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 06-MAY-1999; 99US-0132487.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134221.  
XX 14-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134370.  
XX 18-MAY-1999; 99US-0134768.  
XX 19-MAY-1999; 99US-0134941.  
XX 20-MAY-1999; 99US-0135124.  
XX 21-MAY-1999; 99US-0135353.  
XX 24-MAY-1999; 99US-0135629.  
XX 25-MAY-1999; 99US-0136021.  
XX 27-MAY-1999; 99US-0136392.  
XX 28-MAY-1999; 99US-0136782.  
XX 01-JUN-1999; 99US-0137222.  
XX 03-JUN-1999; 99US-0137528.  
XX 04-JUN-1999; 99US-0137502.  
XX 07-JUN-1999; 99US-0137724.  
XX 08-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 10-JUN-1999; 99US-0138847.  
XX 14-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 16-JUN-1999; 99US-0139453.  
XX 17-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139434.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.  
XX 18-JUN-1999; 99US-0139460.  
XX 18-JUN-1999; 99US-0139461.  
XX 18-JUN-1999; 99US-0139462.  
XX 18-JUN-1999; 99US-0139463.  
XX 18-JUN-1999; 99US-0139750.  
XX 18-JUN-1999; 99US-0139763.  
XX 21-JUN-1999; 99US-0139817.  
XX 22-JUN-1999; 99US-0139899.  
XX 23-JUN-1999; 99US-0140353.  
XX 23-JUN-1999; 99US-0140354.  
XX 24-JUN-1999; 99US-0140695.  
XX 28-JUN-1999; 99US-0140823.  
XX 29-JUN-1999; 99US-0140991.  
XX 30-JUN-1999; 99US-0141287.  
XX 01-JUL-1999; 99US-0141842.  
XX 01-JUL-1999; 99US-0142154.  
XX 02-JUL-1999; 99US-0142055.

```

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143297.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145916.
PR 27-JUL-1999; 99US-0145919.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.

```

```

PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

## Query Match

Best Local Similarity

Matches 210; Conservative

41.9%

Score 978.5;

DB 21;

Length 491;

Pred. No. 3.2e-78;

Indels 41;

Gaps 8;

```

QY 12 LLLPSLSLLFLILKRRNRKTRPNLPGRSGWPFAGETGYLKPATATLGDPMOCH 71
DB 26 LLLLSiaagfl---lllrrtyrmgjlppgsjglpljsetqljgayktenpepfider 82
QY 72 VSKYKIRYRNLFGEPPIVSADAGLNRFILONEGRLEFECSTYRSIGIGLIGKWSMLVGD 131
DB 83 varysvfmthlfgpeplfsadpelnrfvlqnegklfecspaslonllgkhslllmkgs 142
QY 132 MRDMRSISLNLFLSHARLRTILKQVERITLVLDSWOONSIFSAODEKAKKTFPNIMAKH 191
DB 143 lkrhmslmsfianssilkdhmlidrtvrlndswssrvll--meeakkitfcltvqk 200
QY 192 IMSMDPGEETPOLKKEYTTFPMKGVVSAPLNLPGTAVHKLQSRATILKFRKKMERKL 251
DB 201 lmsfdpg-ewseslirkeyllviegffsblpifstlytraiaqr-----rkvaeatl 251
QY 252 DIKEEDQEEVEYKTEDEAEMSKSDHVAKRORTDDDLGLWYLNKHSNLTSTEOILDLISLFA 311
DB 252 vvmvkrreeeeegae-----rkkdmiaallaadddgfdeefvdlvalliava 297
QY 312 GHETSVALALAIIFLOCPKRAVEELREHNLIAARKKLGSELNMDYKKMDFTQCVI 371
DB 298 gyetstlmrlavkflteplalaqlkeehkrlrmksd--syslswsdyksmpfqqcgv 355
QY 372 NETLRGNVVRFLHRKALKDVRKYKGDIPIGSGMKVLPVIAVHLDNSRYDOPMLFNPWMQ 431
DB 356 netlrvaanlgyvrramdvvelkykikpkykvsfsvavhldpnhfdattfprwvq 415
QY 432 QONNGASSSGSGSFSTWGN--YMPREGGPRLCAGSELAKLEMAVFLHHLVLFKNMELAD 490

```

Db 416 -----snsvtspnsufpfgggprlcpgyelarvalsvflhrvtgfswpaeq 465  
Qy 491 DDPFAPFVDFPNCLPIRVR 511  
|: | | | |  
Db 466 dkivffptttrtkryplfvkr 486

RESULT 6  
ID AAG44570 standard; Protein; 492 AA.  
AC AAG44570:  
XX  
XX 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55846.  
DE  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
OS  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140355.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.



```
PR 13-AUG-1999; 99US-0148565-
PR 13-AUG-1999; 99US-0148684-
PR 16-AUG-1999; 99US-0149368-
PR 17-AUG-1999; 99US-0149175-
PR 18-AUG-1999; 99US-0149426-
PR 20-AUG-1999; 99US-0149722-
PR 20-AUG-1999; 99US-0149723-
PR 20-AUG-1999; 99US-0149929-
PR 23-AUG-1999; 99US-0149902-
PR 23-AUG-1999; 99US-0149930-
PR 25-AUG-1999; 99US-0150566-
PR 26-AUG-1999; 99US-0150884-
PR 27-AUG-1999; 99US-0151065-
PR 27-AUG-1999; 99US-0151066-
PR 27-AUG-1999; 99US-0151080-
PR 30-AUG-1999; 99US-0151303-
PR 31-AUG-1999; 99US-0151438-
PR 01-SEP-1999; 99US-0151930-
PR 07-SEP-1999; 99US-0152363-
PR 10-SEP-1999; 99US-0153070-
PR 13-SEP-1999; 99US-0153758-
PR 15-SEP-1999; 99US-0154018-
PR 16-SEP-1999; 99US-0154039-
PR 20-SEP-1999; 99US-0154779-
PR 22-SEP-1999; 99US-0155139-
PR 23-SEP-1999; 99US-0155486-
PR 24-SEP-1999; 99US-0155659-
PR 28-SEP-1999; 99US-0156458-
PR 29-SEP-1999; 99US-0156596-
PR 04-OCT-1999; 99US-0157117-
PR 05-OCT-1999; 99US-0157753-
PR 06-OCT-1999; 99US-0157865-
PR 07-OCT-1999; 99US-0158029-
PR 08-OCT-1999; 99US-0158232-
PR 12-OCT-1999; 99US-0158369-
PR 13-OCT-1999; 99US-0159293-
PR 13-OCT-1999; 99US-0159294-
PR 13-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159329-
PR 14-OCT-1999; 99US-0159330-
PR 14-OCT-1999; 99US-0159331-
PR 14-OCT-1999; 99US-0159637-
PR 14-OCT-1999; 99US-0159638-
PR 18-OCT-1999; 99US-0159584-
PR 21-OCT-1999; 99US-0160741-
PR 21-OCT-1999; 99US-0160767-
PR 21-OCT-1999; 99US-0160768-
PR 21-OCT-1999; 99US-0160770-
PR 21-OCT-1999; 99US-0160814-
PR 21-OCT-1999; 99US-0160815-
PR 22-OCT-1999; 99US-0160980-
PR 22-OCT-1999; 99US-0160981-
PR 22-OCT-1999; 99US-0160989-
PR 25-OCT-1999; 99US-0161404-
PR 25-OCT-1999; 99US-0161405-
PR 25-OCT-1999; 99US-0161406-
PR 26-OCT-1999; 99US-0161359-
PR 26-OCT-1999; 99US-0161360-
PR 26-OCT-1999; 99US-0161361-
PR 28-OCT-1999; 99US-0161920-
PR 28-OCT-1999; 99US-0161921-
PR 28-OCT-1999; 99US-0161993-
PR 29-OCT-1999; 99US-0162142-
PR 29-OCT-1999; 99US-0162142-

Query Match 36.5%; Score 978.5; DB 21; Length 492;
Best Local Similarity 41.9%; Pred. No. 3,2e-78;
Matches 210; Conservative 79; Mismatches 171; Indels 41; Gaps 8;
```

```
DB 84 varysvfmthlfgpeptlfsadpncrnfylngcklfeesypasclnllgkshlllmgs 143
OY 132 MHRDMRSISLNFSLHARLRTILKDVERRHTLEVLDSMOONSIFSADOKKKFTFNLMAKH 191
DB 144 lkhmrhslmsfanssllkhdldidrtvrfnlawsrsvll--meakkltfelvkg 201
OY 192 IMSMDPGEETEQLKKEVYTFMKGVVSAPLNPGRAYHKAOSRATILKFERKMEERKL 251
DB 202 lmsfdpg-ewseslkeylviegflsfipflsfstlyrkaiaqar-----rkvaealt 252
OY 252 DIKEEDQEEEEVKTDEDEMSKSDHVRKORTDDDLGLWLNKSNLSTEQILDLILSLFLFA 311
DB 253 vvmkrrreeeegae-----rkkdmiaallaaddgtsdeeiavfvaliva 298
OY 312 GHETSSVALAIALIFPLQACPKAVEELREHLEIARAKKELGSEINMDYKMDTQCVY 371
DB 299 gyetstlmclavkfltecpplaqkkeehekframksd--syslwsdyksmptfgcv 356
OY 372 NETLRLGNVVRFLHRRKADVRYKGVDPISGKVLPIYSAVHLDSRVDQPLFNPWRMO 431
DB 357 netlrvaanllgsvfrramdvakgklykypkvwvssfravhldpnhkdartfnpwtq 416
OY 432 QONNGASSSGSGSFTWGN--YMPFGGCPRLCAGSELAKLENAVFIHLVLKFNMLAED 490
DB 417 -----snsvctgpnsvftpgfsgprlcpgyelarvalsvflhrlvtgfsvpaeg 466
OY 491 DQFAPPEYDPPNGLPIRYSR 511
DB 467 dkvlfftrtqtgkryplfvr 487
```

## RESULT 7

AAG44572 standard; Protein; 444 AA.

AAG44572;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55848.

KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.

OS Arabidopsis thaliana.

EP103405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.





```
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145145.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 35.38; Score 946; DB 21; Length 444;
Best Local Similarity 41.88; Pred. No. 2.1e-75;
Matches 198; Conservative 77; Mismatches 161; Indels 38; Gaps 7;

QY 39 LPPCKSGMPFICETIGVTKPYTATTLDGPMOQHSKYGKIRSNLFEETPVTSADAGLN 98
DB 3 LPPGSLDPLIGELFGLIGAYKTEPPEFIDETAVATYGVMTNLDIGELFISADPETNR 62
QY 99 FIDNEGRLEFECYSYRSIGGLGKMSMLVLDGNHMRISLSNLFSHARLTILKDV 158
DB 63 FVLQNEGKLFECYSYRSIGGLGKMSMLVLDGNHMRISLSNLFSHARLTILKDV 122
QY 159 RHTLFVLDSMOQNSIFSAODEAKFTFNMAKKNMSMDPGEETEQUKKEYVTFMKGV 218
DB 123 RLVFNLDVSSRVLL--meeakktfclvfkqmsfdpg--ewesaelrkeyllvlegffs 179
QY 219 APULNPGTAYHAKALOSRATILKFERKMEERKLDIKEDDEEVEFKTEDEAKMSKDHVR 278
DB 180 LPLPLFTLYTKALGAT-----rkvaallvymkkrteeeegge----- 219
QY 279 KORTDDLLGWLKNSLNSTEQIIDLILSLFAGHETSSVAIALAIFLQACPKAVEELR 338
DB 220 --rkkdmllaalaaaddgfsdeeltvdlviallvagyetstlmtlavkflletrplalaqk 277
QY 339 EENHETIRAKKELGESELMNDVYKKMDFQCVINETRLGNVVFLLRKALKOVRKYGYD 398
DB 278 eehetkiramksd--sylewsdyksmpftgcynveltrvanllggvfrtamtdevelkygk 335
QY 399 IPSGKVLPIYSAVHLNDSNRDQPNLNPWRMOQNNQASSGSGSFTWGN--YMPFGG 457
DB 336 ipgkwkvfssfrvhdlnpfnkdeartfnprwg-----snvftgspnsvftrpgg 385
QY 458 GPRLCAGSELAKLEMAVFIHLLVLFKFWMLAEDDQPAFPFVDFPNCGLPIRVSR 511
DB 386 gprlcpgrtelarvalsvflhrlvtgfswpaeqdklvftrtqtqkryplfvkr 439
```

RESULT 9  
AAG46490  
ID AAG46490 standard; Protein: 461 AA.  
XX  
AC AAG46490;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58494.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PC 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125768.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144086.  
PR 16-JUL-1999; 99US-0144088.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144816.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.

PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-015138.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154033.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158366.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

```

OY 186 ----NLAAKIMMSDDPEEETEOUKRYEYFMKGVSAPNLPGTQYHNALOSRAATLK 240
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 176 tqIagNlKkPv-----eeefkafeIvIvgvltspIdpIgtvtyrogIqamIdr 223
OY 241 FIERMERKALDIKEEDOEBEKVKTEDBAEMKSDHVKORKOTDDLLGWLKHSN---L 236
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 226 IlreImJerr-----dsge-----tfcdmJgyImkKeggnrYpI 258
OY 297 STEQILDIILISLFAGHETSSVAIALAIFLOACPRAVEELREEBELIAAKKELGESEL 356
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 tdeefrdqvtvltIysyvetvstammaIkYhIdhpkaIdqelaeNaIafreKtq--depl 316
OY 357 NMWDYKKMDFQCVINETLRLGNVYRFLHRKALADVKKYGDIPSGKAVLPVISAHLDN 416
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 gIedvKsmKftravIyestriaIvgtIctIdaIngIvIpkvyrIyvvtreInyda 376
OY 417 SRYDOPNLFEPRWQOONNGASSSGSSGSSFTWNGNNYMPFGGPIRLGSEIALEMAVFI 476
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 nlyedpIIfnpwrwmkksIesq-----nscIvfggvtIrlcpYkgeIvetsfI 425
OY 477 HHVLKFMWELAEDDOPAPAFVDFPGLRPIRS 510
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 426 hYvtvryweeIggdeImvIprvtarPkgvInItIs 459

```

RESULT	10
AAG6489	
ID	AAC6489 standard; Protein: 462 AA.
XX	
AC	AAC6489;
DT	18-OCT-2000 (first entry)
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 58493.
XX	
KM	Protein identification: signal transduction pathway; metabolic pathway;
XX	Hydrolisation assay; genetic mapping; gene expression control; promoter
XX	termination sequence.
OS	Arabidopsis thaliana.
PN	EPI033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125768.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129645.
PR	19-APR-1999; 99US-0130077.
PR	21-APR-1999; 99US-0130449.
PR	23-APR-1999; 99US-0130510.
PR	28-APR-1999; 99US-0130891.
PR	30-APR-1999; 99US-0131449.
PR	30-APR-1999; 99US-0132048.
PR	04-MAY-1999; 99US-0132407.
PR	05-MAY-1999; 99US-0132484.
PR	06-MAY-1999; 99US-0132485.
PR	06-MAY-1999; 99US-0132486.
PR	07-MAY-1999; 99US-0132487.
PR	11-MAY-1999; 99US-0132863.
PR	14-MAY-1999; 99US-0134256.
PR	14-MAY-1999; 99US-0134218.
PR	14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 28-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149368.  
PR 18-AUG-1999; 99US-0149175.  
PR 20-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151080.  
PR 31-AUG-1999; 99US-0151303.  
PR 01-SEP-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153738.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 13-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 18-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.





```
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159684.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 27.1%; Score 726; DB 21; Length 465;
Best Local Similarity 32.3%; Pred. No. 6,7e-56;
Matches 166; Conservative 91; Mismatches 181; Indels 76; Gaps 10;

OY 12 LILPDLILFLILFLILKRRN--KTRFNLPKSGMPFLGETIGYLRPYATTLIGDMQ 69
   ||: || || || || || || || || || || || || || || || || || || ||
DB 11 llllvslcall-----rwmqmytknglppytmgwplfgteteftlkqgp-----nfmr 59

OY 70 QHVSXGKIYRSNLFGEPTIYASDAGLNRLFLQNGRFLFECSYPSIGIGIKSMVLIV 129
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 60 ngrlrygsfxfshllgcplllsmdevrlyllkneskglvpyqpsmldllytcmaavh 119

OY 130 GDMHRDMRSISLNFSLHARLFTILKDYERHTLFLVDSMOQNSIFSADDEAKKFTF---- 185
   | || | || | || | || | || | || | || | || | || | || | || | || |
DB 120 gshrlmrsgslslstsscmrhdllpkvdhmrslldqmelevldldkckhmaflasl 179

OY 186 -----NLAKHIMSDPBEETEOUKKEVYFMKGVNSAPLNLPGTAVNHKALQSRATLIK 240
   || : || : || : || : || : || : || : || : || : || : || : || : ||
DB 180 tqiaqnllrkpfv-----eeftkafiklvvgclsvpldplpctnyrcyqigatnmdr 229

OY 241 FIERKMEERKLDIKEDDEEEEVKTEDEAEMSKSDHVKKORPDDLLGMVLKHSN----L 296
   || || || || || || || || || || || || || || || || || || || || ||
DB 230 llrelmgerr-----dsge-----ftcmldylnmkegnrypl 262

OY 297 STEQILDILSLFLFAGHETSSVALALAFILOACPKAVEELREEHILETARAKKEGESEL 356
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
DB 263 tdeetrdqvvtllysgyetvstsmmalikyldhpkalqelnaehlafrerkrq--depl 320

OY 357 NMDDYKKMDFQOCVINEFLRLGNVVRFLHRRALKVYRKGYIRPGGKVLVPLYSVHNLDN 416
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 321 gledvksmkftravlyetstlativngvllfkttrldenglylbpkxwilyvyltreindya 380

OY 417 SRYDQNLFNPRWQOONNAGSSSGSFSFTWGNVMPFGGPRLCAGSELAKLEMAVFI 476
   :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 361 nlyedpdlfnprwmkkslesg-----nscfvygggtrtlcpkgeiglyveissfi 429

OY 477 HHVLKFMWELAEDDQPAFPVDEPNGLPIRVS 510
   | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

DB 430 hyfctiryweeioggdelmvpfvrkfyghlrlis 463  
RESULT 12  
AAG1836  
ID AAG1836 standard; Protein; 461 AA.  
XX AAG1836;  
-XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10708.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139809.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142390.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142820.  
PR 13-JUL-1999; 99US-0142877.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148365.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159285.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 26.6%; Score 712; DB 21; Length 461;  
Best Local Similarity 32.5%; Pred. No. 1; ie-54;

Matches 164; Conservative 93; Mismatches 196; Indels 52; Gaps 8;

OY 12 LLLPSSLILLFLILKRRN-RKTRFNLPKSGMPLFAGETGYLKPYATATLGDPMQ 69  
Db 1 mmilglilvltcltalrrwgmymyskglppgtmwpifgeteflkgp-----dmk 55  
OY 70 QHVKSKYKTIYNSNFGERTIYASADAGLNRFLLQNEGRLLFECSYRISIGILGKMSMLVLY 129  
Db 56 nqrllyrsfifshilgpcilvsmdaelnryllmneskylvagypqsmldlgctcnlaavh 115  
OY 130 GDMHRDRSISLNLNLSHARLTITLLKOVERHTLFLVDSWQONSIFSADAKKFTFNMA 189

Db 116 gpsnrlmrgslslslsptmmchlllpkldfmmrylvgwddetevdidektkmaf-iss 174  
OY 190 KHMSMDGEEETEQLKKEYVTFMKGVSADPLNPGTAHYKALOSRATILKFIERKMEER 249  
Db 175 llqiaetlkpeveeyrteffekflvgctlsyplldpgtnrysgvgarmidrltlelmgqr 234  
OY 250 KLDIKEEQEEVEYKTEDEAEKMSKSDHVRKORTDDLLGWVLKSN-----LSTEQILDLI 305  
Db 235 k-----esgeft-----dmgylmkdednylltdkeirdqv 267  
OY 306 LSLFAGHETSSVALIALAIFLQACPKAVEELREBEETIAAKKELGSELSLMDYKMD 365  
Db 268 vtllysgyeltvstcsmmaikyldhpkakeelrethlairekrp--depltdidksmk 325  
OY 366 FTQCVINETLRNGVNFRLHAKALDVRKYGVDIPSGKVLVPIASVHLDSRYDQPNLF 425  
Db 326 ftravifetsrlactlvgvltcthdlelmgylldpgvrylvytrelngydslyedpmif 385  
OY 426 NPMRWQOQNNNGASSSGSFSFTWGNMYPFGGPRLCAGSELAKLEMAVFIHLVLRNM 485  
Db 386 npwrwmekslesks-----yflldggvrlcpqkelglssevsfihyftvkyrw 434  
OY 486 ELAEDDQFAPFPVDFPNGLRIRWS 510  
Db 435 eengedkrmvfprvsapkyhlcas 459

RESULT 13  
AAG11835

ID AAG11835 standard; Protein: 462 AA.

AC AAG11835;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 10707.

KW Protein identification: signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PE 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130049.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138647.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139465.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 23-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151338.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 9905-0160899.  
PR 25-OCT-1999; 9905-0161404.  
PR 25-OCT-1999; 9905-0161405.  
PR 25-OCT-1999; 9905-0161406.  
PR 26-OCT-1999; 9905-0161359.  
PR 26-OCT-1999; 9905-0161360.  
PR 26-OCT-1999; 9905-0161361.  
PR 28-OCT-1999; 9905-0161920.  
PR 28-OCT-1999; 9905-0161992.  
PR 28-OCT-1999; 9905-0161993.  
PR 29-OCT-1999; 9905-0162142.

## Query Match

26.6%; Score 712; DB 21; Length 462;

Best Local Similarity 32.5%; Pred. No. 1,1e-54;

Matches 164; Conservative 93; Mismatches 196; Indels 52; Gaps 8;

QY 12 LLLPSLLSLFLILKRRN--KTRRNLPKSGKMPFLGCTGYLKYPTATTLGDMQ 69  
DB 2 mmllgllivlvccltalrlwqmqmryskg1ppgtmgwplfgeteflkqgp----dfmk 56  
QY 70 QHVSQKTYRNSNLFGEPTIVSADAGLNRFITONEGRFLFESQPSRISGIGLKMSMLVLY 129  
DB 57 nqrlrygsffshlqcpvlvsmdaelnrylimeskylvagypqsmldlqtcniaavh 116  
QY 130 GDMHRDRSISLNFSLHARLTLLKDVERTLFLDSWOONSIFSQDEAKKFTFNMA 189  
DB 117 gprhrlmgslsls1pmtmckdhlpkldfmrnylqgwdelevdqeckkmaaf-1ss 175  
QY 190 KHMSMDGEETEQLKKEYTYTFMKGVASAPLNPGTAHKALOSRAITLFTIERKKEER 249  
DB 176 llqaelkpkveeyetefklyvgtlsvpidipgtnyrsqvarnidrllelmgqr 235  
QY 250 KLDIKEDQEEVKTDEAMSKSDHYRKQRTDDLLGWLKHSN---LSTEQIIDL 305  
DB 236 k-----esgeft-----dmgylmkrednrylltckelrtqv 268  
QY 306 LSLFAGHETSSVALAIAFLIACPRAVEELREHELEIARAKKELGSELNWDYKMD 365  
DB 269 vtllysgvetsstmmakylhdpkaleelrhehlairekrp--depltdldiksmk 326  
QY 366 FTQGVINETLGNVNFRLHAKLDVRYKGYDIPSGKVLPIVISAHLDSRRDQMLF 425  
DB 327 ftravifetsrlativgylrktchdlnngyl1pkgyrlyvvtrelhdslyedpmif 386  
QY 426 NPMRMQOONNGASSSGSFSFTWGNMNPFGGPRLCAGSLALEMVFTHHIVLKFNW 485  
DB 387 npwtmekslesks-----ylllgygyrlcpqkelgisevssllhylvlkytw 435  
QY 486 ELAEDDQFAFPFVDFPNGLPYRS 510  
DB 436 eengedklmvfprvsapkgyhlks 460

## RESULT 14

ID AAG11834 standard; Protein; 465 AA.

XX AAG11834.

XX AC AAG11834.

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 10706.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 9905-0121825.  
PR 05-MAR-1999; 9905-0123180.  
PR 09-MAR-1999; 9905-0123580.  
PR 23-MAR-1999; 9905-0125788.  
PR 25-MAR-1999; 9905-0126264.  
PR 29-MAR-1999; 9905-0126785.  
PR 01-APR-1999; 9905-0127462.  
PR 06-APR-1999; 9905-0128234.  
PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129845.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-0130449.  
PR 23-APR-1999; 9905-0130510.  
PR 23-APR-1999; 9905-0130891.  
PR 28-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 30-APR-1999; 9905-0132407.  
PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 06-MAY-1999; 9905-0132487.  
PR 07-MAY-1999; 9905-0132863.  
PR 11-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 18-MAY-1999; 9905-0134378.  
PR 19-MAY-1999; 9905-0134941.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136392.  
PR 28-MAY-1999; 9905-0136782.  
PR 01-JUN-1999; 9905-0137222.  
PR 03-JUN-1999; 9905-0137528.  
PR 04-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139453.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139452.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 22-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140354.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141842.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.





```
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149920.
PR 22-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150084.
PR 26-AUG-1999; 99US-0150086.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 24.4%; Score 653.5; DB 21; Length 468;  
Best Local Similarity 30.7%; Pred. No. 1.8e-49;  
Matches 160; Conservative 91; Mismatches 189; Indels 81; Gaps 13;

```
OY 9 LLELLLPILSLILPILILKRRNRKTRFNLPPGKSGMPPFLGETIGLKPYTATTLGDFM 68
Db 8 vvpil---llcillyvyskkkknstgkllp9smgwpylget---lqlysqnp-nvff 59
OY 69 OQHVSKGYKTYRNSLFGEPITVSADAGLNRFITONEGRLEFECSPRSIGIILKQWSMLVL 128
Db 60 tsbkqkygeifktrilgygpcvmlaspeaarfvltvtnahmktprskelkgpsalfth 119
```

```
OY 129 VGDHHRDMRSISLNFSLHARTLTLKQVERHTLFVLDSWOONSIFSADDEAKKTFNL- 187
Db 120 qgdvhsnrlklyqssfypecitk-lldptehialssqswamprpyslygemkktfdvg 178
OY 188 ---MAKHMSMDPEEETEDLQKKEYTFMKGVVSAPLNLPGTMYHKALOSRATILAFIER 244
Db 179 lalfighles----sykeilkhnynlvdkgynsfmslpqtsyhxkalmarqklktivse 233
OY 245 KMEERKLDIKEEDQEEBEVTEDEBAEMSKSDHVRKORTDDDLGWLKHSN-----LSTE 299
Db 234 licerr-----ekralqtdfighlllnfnkqrvltge 266
OY 300 QIDDLILSLFAGHETSSVALAIFELQACPRAVEELREE---HLEIARAKKEGESE 355
Db 267 qiadnllgvlfaagdttaascitwlkylhddqkllaevekqakalyeansrek-----p 321
OY 356 LNMDDYKKMDFTQCVINETRLGNVVRFLHRKALKDYRYKGYDIPIGSKVLPYISAVHLD 415
Db 322 ltwrqtrmpplthkviveslrmasslstrfaevdvdeykylllpkykwmpilfrnlhna 381
OY 416 NSRYDOPNLFNPRMOQOQNNGASSSGSGSESTWGNMYPGSGPRLCAGSELAKEMAVF 475
Db 382 pkfysnpevidpsrfe-----vnpkpnctimpfgsvhacpynelaklqllif 428
OY 476 IHHLVLKFNNELAEDDO-----PFAPEYDFPNGLPPIRSR 511
Db 429 lnhlvsnfwevkggekgyldpfpdp---gnqlpatftr 465
```

Search completed: September 28, 2001, 18:44:18  
Job time: 48 sec